

Distance-based Exponential Probability Models on Constrained Combinatorial Optimization Problems

Josu Ceberio
University of the Basque Country
UPV/EHU
Donostia-San Sebastian, Spain
josu.ceberio@ehu.eus

Alexander Mendiburu
University of the Basque Country
UPV/EHU
Donostia-San Sebastian, Spain
alexander.mendiburu@ehu.eus

Jose A. Lozano
Basque Center for Applied
Mathematics (BCAM)
Bilbao, Spain
ja.lozano@ehu.eus

ABSTRACT

Estimation of distribution algorithms have already demonstrated their utility when solving a broad range of combinatorial problems. However, there is still room for methodological improvements when approaching constrained type problems. The great majority of works in the literature implement external repairing or penalty schemes, or use ad-hoc sampling methods in order to avoid unfeasible solutions. In this work, we present a new way to develop EDAs for this type of problems by implementing distance-based exponential probability models defined exclusively on the set of feasible solutions. In order to illustrate this procedure, we take the 2-partition balanced Graph Partitioning Problem as a case of study, and design efficient learning and sampling methods in order to use these distance-based probability models in EDAs.

CCS CONCEPTS

• **Mathematics of computing** → **Combinatorial optimization; Probabilistic algorithms; Maximum likelihood estimation;**

KEYWORDS

Constraint, estimation of distribution algorithm, distance-based exponential model, Graph Partitioning Problem

ACM Reference Format:

Josu Ceberio, Alexander Mendiburu, and Jose A. Lozano. 2018. Distance-based Exponential Probability Models on Constrained Combinatorial Optimization Problems. In *GECCO '18: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3205651.3205659>

1 INTRODUCTION

Estimation of Distribution Algorithms (EDAs) [4] have proved to be a powerful evolutionary algorithm for solving either artificial or real-world Combinatorial Optimization Problems (COP). Among different types of COPs, *constrained problems* [3] (problems with restrictions) have shown to be very challenging for EDAs. In general, a n dimensional constrained COP consists of

$$\begin{aligned} &\text{minimizing} && f(\mathbf{x}), && \mathbf{x} = (x_1, \dots, x_n) \\ &\text{subject to,} && g_i(\mathbf{x}) \leq 0, && i = 1, \dots, r \\ &&& h_j(\mathbf{x}) = 0, && j = r + 1, \dots, m \end{aligned}$$

where f is the objective function, \mathbf{x} is a solution of the search space Ω , and g and h are, respectively, inequality and equality constraints.

Constrained problems introduce a challenging characteristic in the definition of the solutions in the search space: feasibility. A solution \mathbf{x} is *feasible* if it holds all the constraints; otherwise it is *unfeasible*.

It must be noted that the search space Ω is usually composed of all the solutions that fit in the codification. This point represents a serious drawback for EDAs, since these algorithms learn a probability distribution defined on the whole set of solutions induced by the codification (either feasible or unfeasible), and thus, when sampling new solutions, both types could be generated. In this sense, in order to hold the feasibility of the solutions, a number of approaches have been proposed [6]. However, in all these cases, the behavior of EDAs is somehow denaturalized as the obtained sample of solutions does not follow the estimated probability distribution.

In this work we address constrained problems in the framework of EDAs following a more general research line: design and implement new distance-based exponential probability models that define a probability distribution only on the set of feasible solutions [2]. In order to design such types of models, we address three key aspects related to the model: (1) calculate the probability of any x in Ω , (2) given a set $\{\mathbf{x}^1, \dots, \mathbf{x}^n\}$ of solutions, estimate the parameters of the probability model, and (3) sample solutions from the estimated model. For the sake of illustrating this procedure, we considered the Graph Partitioning Problem (GPP) [5].

2 2-PARTITION BALANCED GPP

Given a weighted undirected graph $G = (V, E)$ with $n = |V|$ vertices, and the weights w_{ij} between each pair of vertices in G , the balanced 2-partition Graph Partitioning Problem (GPP) is the problem of finding a 2-partition of the vertices in G while minimizing the total weights of the edges between sets [5]. The objective function is formalized as

$$f(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n x_i(1 - x_j)w_{ij}$$

where any solution $\mathbf{x} \in \Omega$ is encoded as a binary vector and x_i indicates the set to which vertex i is assigned. A feasible solution is subject to the constraint of having the same number of zeros and ones, i.e., $\sum_{i=1}^n x_i = n/2$. Taking this into account, the search space of solutions Ω is composed of $\binom{n}{n/2}$ binary vectors.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '18, July 15–19, 2018, Kyoto, Japan

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5764-7/18/07...\$15.00

<https://doi.org/10.1145/3205651.3205659>

3 DISTANCE-BASED EXPONENTIAL MODEL

Let us consider a distance-based exponential probability model P defined on a finite (or infinite countable) domain of solutions Ω . Under this model, the probability value of every solution in the domain, $x \in \Omega$, is calculated as

$$P(x) = \frac{e^{-\theta d(x, \bar{x})}}{\psi(\theta)}$$

where $\psi(\theta)$ is the normalization constant, θ denotes the spread of the distribution, and $d(x, \bar{x})$ is the distance of x to a reference solution \bar{x} . This model assigns every solution x a probability that decays exponentially with respect to its distance to \bar{x} .

For the sake of illustrating the approach presented in this manuscript, we considered the GPP as a case of study, and we will individually address each of the three key aspects described above.

The first decision to make is the election of the distance-metric d , which directly depends on the codification used. As described in Sec. 2, the solutions of the 2-balanced GPP are codified as binary vectors of size n , with an equal number of zeros as ones. Among binary vectors, the most natural distance-metric is the Hamming distance. However, in this case, it is highly redundant as a solution x and its complementary solution $\neg x$ are equal. Therefore, we present the following distance-metric:

Definition 3.1. Let x be a solution for a 2-balanced GPP problem, and $\neg x$ is the negated or complementary solution of x , then the distance-metric $d(x, \bar{x})$ is calculated as

$$d(x, \bar{x}) = \min\{d_H(x, \bar{x}), d_H(\neg x, \bar{x})\}$$

Under this metric, the maximum distance K to which a solution can be defined is $\lfloor n/2 \rfloor$. The minimum distance is 0 (when x or $\neg x$ are equal to \bar{x}). Solutions can only be at paired distances in order to hold the constraint of the balanced 2-partition case. The number of solutions at distance k (being k even) is $\binom{n/2}{k/2}^2$. Note that each possible solution can be coded with two different binary vectors (one the negation of the other).

3.0.1 Computing the probability value. Given a solution x and the parameters θ and \bar{x} , computing $P(x)$ is given by an efficient computation of the normalization constant $\psi(\theta)$. Considering the number of solutions at each distance, this function can be reformulated as follows:

$$\psi(\theta) = \sum_{l=0}^{K/2} \binom{n/2}{l}^2 e^{-\theta 2l}$$

3.0.2 Learning. Given a set $\mathbf{x} = \{x^1, \dots, x^N\}$ of solutions, the learning step consists of estimating the parameters of the proposed model: \bar{x} and θ . In this work, we decided to estimate the MLE of the parameters. To that end, we calculate the likelihood function below

$$\log L(\theta, \bar{x} | \mathbf{x}) = -\theta N \bar{d} - N \log \left[\sum_{l=0}^{K/2} \binom{n/2}{l}^2 e^{-\theta 2l} \right] \quad (1)$$

where $\bar{d} = \frac{1}{N} \sum_{i=1}^N d(x_i, \bar{x})$. This term is independent to that of θ , so by minimizing $\sum_{i=1}^N d(x_i, \bar{x})$ the likelihood is maximized. Such an estimation is an optimization task itself. Therefore, we propose estimating \bar{x} as the mode solution [1].

In order to find the maximum value of θ that maximizes Eq. 1, we compute its derivative and equate it to 0, obtaining the expression

$$\sum_{l=0}^{K/2} \binom{n/2}{l}^2 (2l - \bar{d}) e^{-\theta 2l} = 0. \quad (2)$$

θ cannot be calculated exactly from Eq. 2 and, thus, we propose using numerical methods, such as Newton-Raphson to estimate it.

3.0.3 Sampling. Once the parameters of the model have been estimated, the next step in EDAs is to sample a set of solutions that follows the previous probability distribution. It is worth to remark that (1) every solution has the same number of solutions at each distance, and (2) in the proposed model, all the solution at distance k from \bar{x} have the same probability.

Taking those considerations into account, first, a distance $k = 2l$ at which generate a solution is sampled. The probability under the proposed model to generate a solution at distance k is

$$P(k) = \sum_{x | d(x, \bar{x})=k} P(x) = \binom{n/2}{k/2}^2 \frac{e^{-\theta k}}{\psi(\theta)}$$

Once k has been decided, we generate uniformly at random one solution among those at distance k from the reference solution \bar{x} . To that end, we choose u.a.r. $k/2$ zeros and $k/2$ ones, and the chosen items are bit-flipped.

4 CONCLUSIONS

In this work, we propose using distance-based exponential probability models defined exclusively on the set of feasible solutions. We took as a case of study the 2-balanced Graph Partitioning Problem, and addressed the three key aspects related to the model: (1) calculate the probability of any x in Ω , (2) given a set $\{x^1, \dots, x^n\}$ of solutions, estimate the parameters of the probability model, and (3) sample solutions from the estimated model.

ACKNOWLEDGMENTS

This work has been partially supported by the Research Groups 2013-2018 (IT-609-13) and Elkartek programs (Basque Government), and TIN2016-78365R and Severo Ochoa Program SEV-2013-0323 (Spanish Ministry of Economy, Industry and Competitiveness). Jose A. Lozano is also supported by BERC 2014-2017.

REFERENCES

- [1] Josu Ceberio, Ekhine Iruozki, Alexander Mendiburu, and Jose A. Lozano. 2014. A Distance-based Ranking Model Estimation of Distribution Algorithm for the Flowshop Scheduling Problem. *IEEE Trans. on Evol. Comp.* 18, 2 (2014), 286–300.
- [2] Ekhine Iruozki. 2014. *Sampling and Learning Distance-based Probability Models for Permutation Spaces*. Ph.D. Dissertation. University of the Basque Country.
- [3] Ali Osman Kusakci and Mehmet Can. 2012. Constrained Optimization with Evolutionary Algorithms: A Comprehensive Review. *Southeast Europe Journal of Soft Computing* (2012).
- [4] Pedro Larrañaga and Jose A. Lozano. 2002. *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Kluwer Academic Publishers.
- [5] Elad Mezuman and Yair Weiss. 2012. Globally Optimizing Graph Partitioning Problems Using Message Passing. In *Proc. 15th AISTATS*. 770–778.
- [6] Carlos Segura, Carlos A Coello Coello, Gara Miranda, and Coromoto León. 2016. Using multi-objective evolutionary algorithms for single-objective constrained and unconstrained optimization. *Annals of Operations Research* 240, 1 (2016), 217–250.