

Approaching the Quadratic Assignment Problem with Kernels of Mallows Models under the Hamming Distance

Etor Arza

BCAM - Basque Center for Applied Mathematics
Bilbao, Spain - Basque Country
earza@bcamath.org

Aritz Pérez

BCAM - Basque Center for Applied Mathematics
Bilbao, Spain - Basque Country
aperez@bcamath.org

Josu Ceberio

University of the Basque Country (UPV/EHU)
Donostia-San Sebastian, Spain
josu.ceberio@ehu.es

Ekhiñe Irurozki

BCAM - Basque Center for Applied Mathematics
Bilbao, Spain - Basque Country
eirurozki@bcamath.org

ABSTRACT

The Quadratic Assignment Problem (QAP) is a specially challenging permutation-based np-hard combinatorial optimization problem, since instances of size $n > 40$ are seldom solved using exact methods. In this sense, many approximate methods have been published to tackle this problem, including Estimation of Distribution Algorithms (EDAs). In particular, EDAs have been used to solve permutation problems by introducing distance based exponential models, such as the Mallows Models. In this paper we approximate the QAP with a Hamming distance based kernels of Mallows Models. Based on the benchmark instances, we have observed that our approach is competitive, reaching the best-known solution in 71% of the tested instances, especially on large instances ($n > 125$), where it is able to outperform state of the art results in 43 out of 288 instances.

CCS CONCEPTS

• **Mathematics of computing** → *Combinatorial optimization; Evolutionary algorithms;*

KEYWORDS

Estimation of Distribution Algorithm, Quadratic Assignment Problem, Evolutionary Algorithm, Hamming distance

ACM Reference Format:

Etor Arza, Josu Ceberio, Aritz Pérez, and Ekhiñe Irurozki. 2019. Approaching the Quadratic Assignment Problem with Kernels of Mallows Models under the Hamming Distance. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion)*, July 13–17, 2019, Prague, Czech Republic. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3319619.3321976>

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic

© 2019 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.
ACM ISBN 978-1-4503-6748-6/19/07...\$15.00
<https://doi.org/10.1145/3319619.3321976>

1 INTRODUCTION

The Quadratic Assignment Problem (QAP) is a well known NP-hard combinatorial permutation-based problem. The problem consists in optimally allocating n facilities at n locations, in order to minimize a cost function related to the flow and the distance between every pair of facilities. The QAP is considered to be a difficult problem. In fact, medium and large non-symmetric instances ($n > 40$, where n is the number of items in each permutation) still remain computationally unfeasible for exact methods. As a result, many metaheuristic methods such as Genetic Algorithms and Simulated Annealing have been applied to the QAP.

Estimation of Distribution Algorithms (EDAs) have also been used to solve permutation based problems by introducing probabilistic models on the space of permutations of n items (S^n) [1]. The MM is an exponential distribution based on a distance in S^n . The EDA - MMs in the literature are based on the Cayley, Ulam and Kendall distances, while the Hamming distance has not previously been used for EDAs. The Hamming distance between two permutations counts the number of point-wise disagreements and is a natural choice for measuring the distance between assignments. Unlike other distances, such as Kendall, the relative ordering between pairs of items in the permutation is not taken into account under the Hamming distance. This is a basic characteristic in assignments or matchings. Moreover, there have been studies in the literature suggesting that an EDA based on a particular distance for a problem will give good results if that distance is a natural measure of disagreement for that problem [3].

Another property of the MM is that it is an unimodal distribution centred at a given central permutation. The unimodality and symmetry properties imposed by the MM can be too restrictive in certain contexts and do not allow the modelling to be flexible [6]. An alternative that breaks these strong assumptions, is the kernel density estimate using Mallows kernels (KMMs). Instead of having a central permutation, KMMs spread the probability mass by using a non-parametric averaging of MMs centred at each data point. Taking advantage of this flexibility and the assignment nature of the QAP, the algorithm proposed in this paper is a Hamming based KMM EDA. The convergence of our algorithm is controlled with the sharpness parameter of the KMM, using a simple exploration-exploitation scheme. In order to enhance the performance of the algorithm, we hybridized the EDA with a local search algorithm.

2 DISTANCE BASED PROBABILITY MODELS

The Mallows Models (MM) is a family of probabilistic models on the space of permutations of size n (S^n). The MM is one of the most simple and natural distributions on S^n , requiring just two parameters: the concentration parameter $\theta \in \mathbb{R}^+$ and the location parameter, $\sigma_0 \in S^n$ [5]. The location parameter, also known as the central permutation, is the mode of the distribution. For the rest of the permutations, the probability of each permutation σ decreases exponentially with respect to its distance from the central permutation $d(\sigma, \sigma_0)$. The speed of this exponential decay is controlled by θ , the sharpness of the distribution.

The MM can be based on different distances. Moreover, the distance used on the EDAs seems to be correlated with the performance of an EDA on a particular problem family, as suggested in [3]. In this paper, we introduced a Hamming based EDA to the framework of permutation-based combinatorial problems.

The unimodality property of the MM can be too limiting in certain situations, and Kernels of Mallows Models have been proposed as an alternative [6]. Given a set of k permutations, the KMM is the non-parametric averaging of k MMs centred in these permutations. Formally, given the set of k central permutations $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_k\}$, the mass probability distribution of KMM can be defined as $K(\sigma|\sigma, \theta) = \frac{1}{k} \sum_{i=1}^k \frac{e^{-\theta d(\sigma, \sigma_i)}}{\psi(\theta)}$ where $\psi(\theta)$ is the normalization constant.

3 HAMMING BASED KMM EDA

Any EDA has three parts: selection, learning and sampling. The selection procedure proposed in this paper includes sorting the objective function values and selecting the best half of the population. In addition, elitism was used on the set of selected permutations.

There is no real learning phase in our approach. Instead, the convergence of the EDA is controlled by a simple exploration-exploitation scheme. The trade-off is balanced by using θ , the concentration parameter of KMMs that increases over the iterations.

Roughly, the sampling procedure is as follows: We center an MM in a permutation chosen u.a.r from the selected set of permutations (where θ is set according to the exploration-exploitation scheme) and sample a new permutation from this distribution. In particular, we use the Distance Sampling Algorithm, see [5] for details and implementations.

As stated previously, our algorithm hybridizes an EDA with a best-first local search procedure. The local search only takes place after the EDA has converged, and takes the solutions found by the EDA into local optima.

4 EXPERIMENTATION

Some popular instances of the QAP have been employed to evaluate the performance of the proposed approach. In particular, the instances considered in this paper were obtained from the QAPLIB, from the Taixxey instances set [4] as well as from a recently published paper [2].

In order to adjust the parameters of the algorithm, we carried out a previous set of experiments. First, we studied the behaviour of the algorithm in a small set of four instances from the Taixxey instances set [4]. We observed that some of the parameters are more time-dependent than others. Therefore, using grid search, we set

Table 1: Summary of results

n	number of instances	reached the best known	surpassed the best known	CPU time interval (s)
[5,20)	67	100.0%	0.0%	[4,12]
[20,40)	91	83.5%	4.4%	[12,33]
[40,75)	46	47.8%	0.0%	[31,77]
[75,125)	39	5.1%	0.0%	[71,194]
[125,175)	24	83.3%	79.2%	[156,540]
175 and 256	21	95.2%	95.2%	[60,1004]

the most performance critical parameters. Then, we estimated the optimal values for the rest using Bayesian Optimization.

The performance of our approach was tested on a total of 288 instances. For each instance, we run the algorithm 9 times, recording the best and the median score, the median number of evaluated solutions and the median CPU time.¹

In 71% out of the 288 instances tested, the state of the art result was reached. Moreover, in 43 instances, a new best solution was found. The summarized results are presented in Table 1.

5 CONCLUSION

In this paper, we proposed an algorithm to solve the QAP with a Hamming distance based Estimation of Distribution Algorithm. We used the Hamming because it respects the assignment nature of the QAP. The use of KMMs in our model also allows a simple yet effective exploration-exploitation procedure to be implemented through the spread parameter θ of the KMMs. To improve the efficiency of the algorithm, we hybridized our approach with a local search procedure. Out of the tested 288 benchmark instances, in 71% of them, the state of the art result was reached. Moreover, for 43 instances, a new best solution was found. The experimentation stated that our hybrid algorithm is a competitive approach to solve the QAP, particularly on large instances ($n \geq 125$).

Acknowledgments

This work is partially supported by Basque Government (Elkartek program) and the Spanish Ministry of Economy and Competitiveness MINECO: the project TIN2016-78365-R, BCAM Severo Ochoa excellence accreditation SVP-2014-068574 and SEV-2013-0323, and through the project TIN2017-82626-R funded by (AEI/FEDER, UE).

REFERENCES

- [1] J. Ceberio. *Solving permutation problems with estimation of distribution algorithms and extensions thereof*. PhD thesis, UPV/EHU, 2014.
- [2] J. Ceberio, B. Calvo, A. Mendiburu, and J. Lozano. Multi-objectivising combinatorial optimisation problems by means of elementary landscape decompositions. *Evolutionary computation*, pages 1–21, 2018.
- [3] J. Ceberio, E. Irurrozki, A. Mendiburu, and J. A. Lozano. A review of distances for the mallows and generalized mallows estimation of distribution algorithms. *Computational Optimization and Applications*, 62(2):545–564, Nov 2015.
- [4] Z. Drezner, P. M. Hahn, and É. D. Taillard. Recent advances for the quadratic assignment problem with special emphasis on instances that are difficult for meta-heuristic methods. *Annals of Operations research*, 139(1):65–94, 2005.
- [5] E. Irurrozki, B. Calvo, and J. A. Lozano. Mallows and generalized mallows model for matchings. *Bernoulli*, 25(2):1160–1188, 05 2019.
- [6] G. Lebanon and Y. Mao. Non-parametric modeling of partially ranked data. *Journal of Machine Learning Research*, 9(Oct):2401–2429, 2008.

¹Supplementary material with all the results as well as source code is available for the interested reader to download at <https://github.com/EtorArza/GECCO2019>