Visualizing Feasible Regions for Optimization Problems on High-Dimensional Permutations using Dimensionality Reduction Methods

Igor Grebennik
System Engineering Department
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
igorgrebennik@gmail.com

Olga Chorna
System Engineering Department
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
olha.chorna@nure.ua

Inna Urniaieva
System Engineering Department
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
inna.urniaieva@nure.ua

***Abstract*—This paper presents an investigation on the usage of modern dimensionality reduction methods for classic combinatorial optimization problems. We propose the use of t-Distributed Stochastic Neighbor Embedding (t-SNE) method to visualize feasible regions on high-dimensional permutations, aiming to avoid the consequences of combinatorial explosion. The results of the study indicate that the proposed approach can provide valuable insights and improve the understanding of the solution space of high-dimensional permutations for the application of local search approaches.**

Keywords—permutations, dimensionality reduction, combinatorial optimization, permutohedron, adjacency, t-SNE method

I. Introduction

Optimization problems on high-dimensional permutations are ubiquitous in various fields such as vehicle routing problems, bioinformatics, cryptography [1-3]. However, the solution space for these problems is often vast, and the combinatorial explosion can pose a significant challenge in finding the optimal solution [4]. In recent years, dimensionality reduction methods have emerged as powerful tools to tackle such high-dimensional problems by mapping the data into a lower-dimensional space while preserving the relevant information [5].

This paper is devoted to the investigation of the usage of modern state-of-the-art dimensionality reduction methods in classic combinatorial optimization problems. The focus of this study is to visualize admissible solution areas for optimization problems on high-dimensional permutations using dimensionality reduction methods, aiming to avoid the consequences of combinatorial explosion. The proposed approach uses t-Distributed Stochastic Neighbor Embedding (t-SNE), a widely used and effective dimensionality reduction technique, to map high-dimensional permutations into a low-dimensional space while preserving their pairwise distances [6]. The resulting visualization can help understand the geometry and topology of the solution space, which is crucial for developing efficient algorithms for solving optimization problems.

The study of dimensionality reduction methods in this area is a promising direction towards increasing the efficiency of solving classic combinatorial optimization problems. The insights gained from the proposed approach can be applied to various fields. The remainder of this paper is organized as follows: in Section 2, we provide a brief overview of combinatorial optimization problems as well as dimensionality reduction methods and their applications. Section 3 presents the problem statement and proposed approach for visualizing admissible solution areas for optimization problems on high-dimensional permutations. In Section 4, we present experimental results on datasets, followed by a discussion of the insights gained from the visualization. Finally, we conclude the paper in with a summary of the contributions and future research directions.

II. Problem Formulation

Combinatorial optimization problems on permutations involve finding the optimal arrangement or ordering of a set of elements or objects [7-9]. These problems have a wide range of applications in various fields. Permutations can be represented as vectors or matrices, where each element or row represents an object, and the order of the elements or rows determines the permutation. The solution space for these problems is often vast and can grow exponentially with the number of objects, making it challenging to find the optimal solution. As a result, developing efficient algorithms for solving combinatorial optimization problems on permutations is a crucial research area.

One of the most well-known combinatorial optimization problems on permutations is the traveling salesman problem (TSP), where the objective is to find the shortest possible route that visits each city exactly once and returns to the starting city. Other examples include the knapsack problem, the job sequencing problem, and the permutation flowshop scheduling problem [10].

Several methods have been developed to solve combinatorial optimization problems on permutations, including heuristic algorithms, exact algorithms, and metaheuristic algorithms [11-13]. Heuristic algorithms, such as greedy algorithms and local search, are simple and fast but may not always find the optimal solution. Exact algorithms, such as branch and bound and dynamic programming, guarantee the optimal solution but may be computationally expensive, especially for large problem instances. Metaheuristic algorithms, such as simulated annealing and genetic algorithms, are a compromise between heuristic and exact methods and can often find high-quality solutions efficiently.

Dimensionality reduction methods have become increasingly popular in various fields due to their ability to transform high-dimensional data into a lower-dimensional space while preserving the relevant information. These methods can be broadly classified into two categories: linear and nonlinear dimensionality reduction [5].

Linear dimensionality reduction methods, such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Canonical Correlation Analysis (CCA), use linear transformations to reduce the dimensionality of the data. These methods are widely used in image and signal processing, bioinformatics, and machine learning.

Nonlinear dimensionality reduction methods, on the other hand, use nonlinear transformations to capture the nonlinear structure of the data. These methods include t-SNE, Isomap, UMAP (Uniform Manifold Approximation and Projection), Locally Linear Embedding (LLE), and Laplacian Eigenmaps. They are particularly useful in visualizing high-dimensional data in a lower-dimensional space while preserving the local structure of the data. Nonlinear dimensionality reduction methods have been successfully applied in various fields, such as natural language processing, bioinformatics, and computer vision [14].

In recent years, machine learning and neural networks have also been applied to combinatorial optimization problems where the objective is to find the optimal solution from a vast solution space of high-dimensional permutations [15-17]. In this paper it is proposed to use subclass of machine learning methods such as dimensionality reduction for mapping the combinatorial data into a lower-dimensional space. This approach can help understand the geometry and topology of the solution space, which is crucial for developing efficient algorithms for solving optimization problems.

III. Admissible Solution Areas On High-Dimensional Permutations

Combinatorics relies heavily on the concept of permutation, which serves as a fundamental cornerstone of the field. The definition of a permutation and of it cyclic structure is given, for example, by R. Stanley, and the sets of permutations by M. Bona. For the purpose of this discussion, we will adopt the following definition of a permutation. [7-8]:

The linear ordering of elements of a single generating set A={a_1, a_2,…, a_n} is called a permutation $\pi = (a_1, a_2,\dots, a_n) = (\pi(a_1), \pi(a_2),\dots, \pi(a_n))$, or, if it is necessary to emphasize the fact that it contains n elements, n– permutation. The set of permutations generated by the elements $a_1 < a_2 < \dots < a_n$ will be writing as P_n.

One of the fundamental properties of permutations is their cyclic structure. Permutation π might be represented as a combination of non-intersecting cycles, writing as $\pi = C_1C_2\dots C_k$ [8].

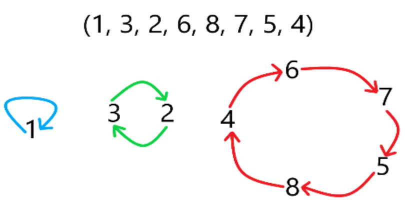

Fig. 1. Random permutation cyclic structure example

Let us consider a combinatorial optimization problem in the following statement:

$$L(p) = \sum_{i=1}^{n} c_i p_i \to \min;$$

$$p = (p_1, p_2, \dots, p_n) \in \bar{P} \subset P_n,$$

where $c_j \in R,\ p_i \in J_n = \{1, 2, \dots, n\},\ i \in J_n\ \ \forall i, j,\ p_i \neq p_j,$

and where $\bar{P}$ is some subset of the permutation set P_n.

As a result of immersion of set $\bar{P}$ into an arithmetic Euclidean space, we can formulate the problem, equivalent to previous problem of optimization of linear function in space R_n:

$$L(x) = \sum_{j=1}^{n} c_j x_j \to \min;$$

$$x = (x_1, x_2, \dots, x_n) \in \bar{E}$$

where $c_j \in R,\ \ j \in J_n$, and $\bar{E}$ is an immersion of the set $\bar{P}$ in Euclidean space.

The problem of minimizing linear functions on Euclidean combinatorial sets of permutations and arrangements without additional constraints is solved, for example, in [8, 17]. But as soon as there appear any sorts of additional constraints on the admissible solution areas the problem is no longer stays trivial. That is why we consider dimensionality reduction techniques that will preserve the local structure of the permutational subsets and make it possible to use different sorts of well-known local search optimization methods.

Another key concept related to permutations is the permutohedron [17, 18]. A permutohedron is a polytope that represents all possible permutations of a set of numbers. In other words, it is the convex hull of all permutations of a given set of numbers. The permutohedron is a fundamental object in combinatorial optimization problems on permutations, and it serves as an admissible solution area for many such problems. In general, the permutohedron serves as a natural admissible solution area for many combinatorial optimization problems on permutations.

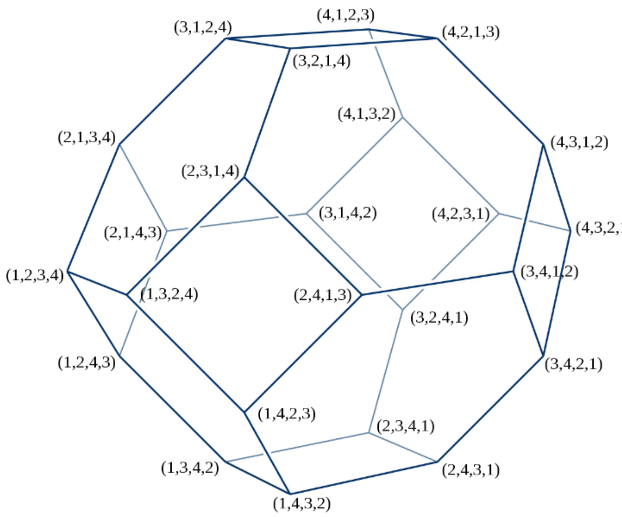

Fig. 2. Permutohedron of order 4, a truncated octahedron

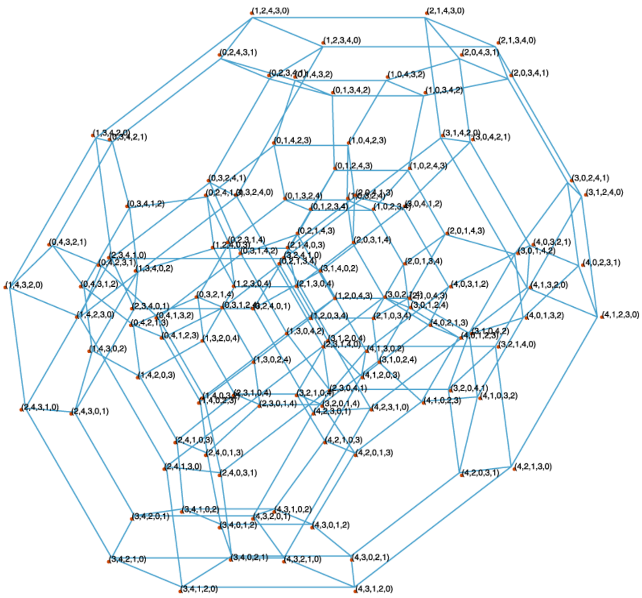

Fig. 3. Permutohedron of order 5, an omnitruncated 5-cell

Visualizing the permutohedron can provide insights into the structure of the solution space, and it can help in developing efficient algorithms for solving these problems. Also visualizing the permutohedron can be challenging, especially in high dimensions. As the number of dimensions increases, the complexity of the polytope increases exponentially, making it difficult to visualize the permutohedron directly. In fact, even for relatively small values of n (e.g., n=10), the number of permutations is already in the order of millions, which makes it practically impossible to visualize the entire permutohedron.

To overcome this problem, dimensionality reduction methods might be used, which can map the high-dimensional permutohedron to a lower-dimensional space while preserving its essential structure. Techniques such as t-SNE and UMAP could be used to visualize the permutohedron in two or three dimensions, which can aid in understanding the structure of the solution space and developing efficient algorithms for solving combinatorial optimization problems on permutations.

However, it is important to note that the lower-dimensional representations of the permutohedron obtained through dimensionality reduction methods are not exact and may distort the structure of the original high-dimensional space. Therefore, it is important to interpret the results of visualization with caution and to validate the findings using other techniques, such as clustering or classification analysis.

In summary, the visualization of the permutohedron on high dimensions is a challenging task, but dimensionality reduction methods can aid in understanding the structure of the solution space and developing efficient algorithms for solving combinatorial optimization problems on permutations. However, it is important to interpret the results with caution and validate them using other techniques.

IV. Dimensionality Reduction On The Set Of Permutations

To investigate the potential of dimensionality reduction methods for permutation data, t-SNE was chosen for its versatility and ease of implementation in the Python library sklearn. A full list of permutations was used as the dataset, with each element in the permutation being treated as a single feature. Additional information on the number of cycles in each permutation's cyclic structure and the adjacency distance to the first permutation (0, 1, 2, …, n) was also included in the analysis. The study aimed to explore the usefulness of t-SNE for visualizing the complex structures of permutation data and provide insights into the nature of permutation optimization problems.

The dataset used in the analysis consisted of permutations with a dimensionality of up to 7, what means that the largest dataset resulting in a total of 5040 permutations with 9 features each.

In the Python programming language, t-SNE is implemented in the scikit-learn library, also known as sklearn. This library provides a range of tools for machine learning and data analysis, including various dimensionality reduction techniques. Moreover, the library also provides a range of options to fine-tune the t-SNE algorithm, such as the perplexity parameter, which controls the balance between local and global structure in the resulting visualization.

As a result of our investigation, we present a 2D graph (Figure 4) that displays the initial point marked with a cross and the points in the permutohedron marked with orange color that are located on different distances such as 1, 2, 3 and 4 transpositions from initial permutation in original permutohedron. This graph provides a visual representation of the adjacency of permutations and their corresponding distances on the permutohedron.

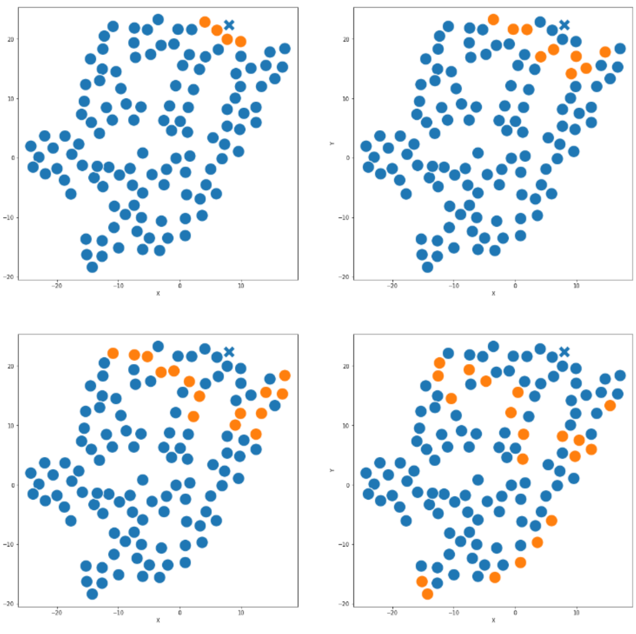

Fig. 4. 2D representation of the adjacency of permutations

During the experiments conducted for this research paper, it was observed that setting the perplexity parameter equal to the total number of adjacent permutations within distances 1 and 2 for each permutation resulted in the most optimal outcomes. The resulting visualization of 5040 permutations of dimensionality 7 using the t-SNE method is shown in the figure below. The heatmap depicts the adjacency of permutations in Euclidean space on the permutohedron.

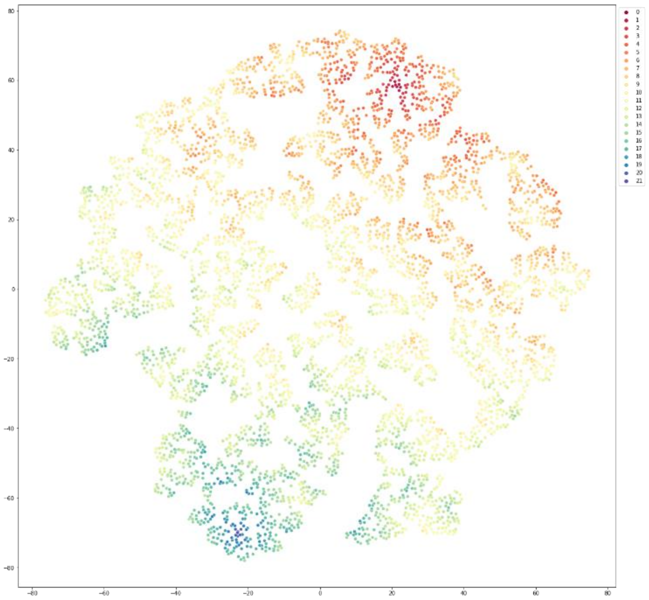

Fig. 5. 2D permutations of order 7 adjacency heatmap

In order to examine the relationship between dimensionality reduction of permutation data obtained through t-SNE and the actual values of combinatorial functions on permutations, 100 linear functions with coefficients randomly generated from the interval [-100, 100] were computed. Subsequently, the visualization of the original values of the linear functions on 2D permutations was obtained.

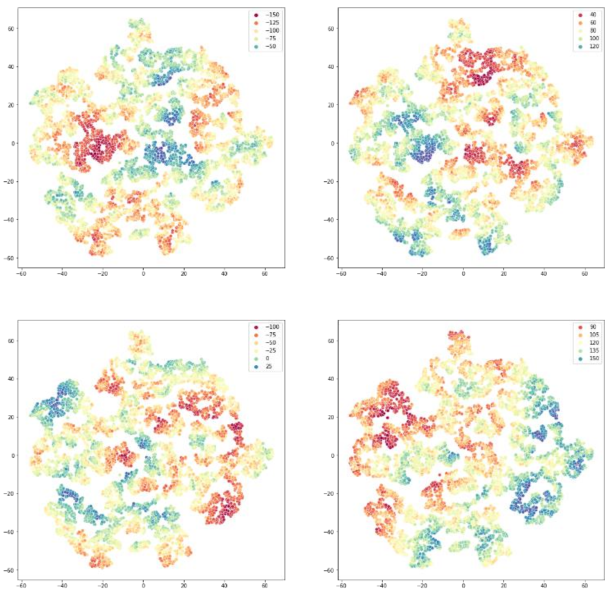

Fig. 6. Linear functions values on 2D permutations heatmap

It should be emphasized that the heatmap of the linear function values on 2D reduced permutations contains areas that correspond to local or global maximums and minimums of different linear functions.

V. Conclusion

In conclusion, this paper presents a novel approach to address the combinatorial explosion problem in classic combinatorial optimization problems. The proposed method uses t-SNE, a modern dimensionality reduction technique, to visualize the admissible solution areas of high-dimensional permutations. The results of the study indicate that the proposed approach can effectively reduce the dimensionality of the problem, and provide valuable insights into the structure of the solution space. Moreover, the study of dimensionality reduction methods in this area is a promising direction towards improving the efficiency of solving classic combinatorial optimization problems. Overall, this research opens up new opportunities for applying modern techniques to traditional combinatorial optimization problems, and can potentially lead to significant advancements in this field.

References

[1] Grebennik, Igor, Olga Chorna, and Inna Urniaieva. "Distribution of Permutations with Different Cyclic Structure in Mathematical Models of Transportation Problems." 2022 12th International Conference on Advanced Computer Information Technologies (ACIT). IEEE, 2022.

[2] R. Dupas, I. Grebennik, I. Litvinchev, T. Romanova, O. Chorna, “Solution strategy for one-to-one pickup and delivery problem using the cyclic transfer approach,” in EAI Endorsed Transactions on Energy Web, vol. 20 (27), pp. 1–9, 2020.

[3] C. Gravel, D. Panario, D. Thomson , “Unicyclic strong permutations,” in Cryptography and Communications, vol. 6(11), pp. 1211–1231, 2019.

[4] Burton A. Leland, Bradley D. Christie, James G. Nourse, David L. Grier, Raymond E. Carhart, Tim Maffett, Steve M. Welford, and Dennis H. Smith “Managing the Combinatorial Explosion” Journal of Chemical Information and Computer Sciences 1997 37 (1), 62-70.

[5] Huang, Xuan, Lei Wu, and Yinsong Ye. "A review on dimensionality reduction techniques." International Journal of Pattern Recognition and Artificial Intelligence 33.10 (2019): 1950017.

[6] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9.11 (2008).

[7] M. Bona, “Combinatorics of permutations,” Chapman & Hall/CRC, 2004.

[8] R.P. Stanley, “Enumerative combinatorics,” Cambridge: Cambridge University Press, 1999.

[9] Ch.A. Charalambides, “Enumerative combinatorics,” Chapman and Hall/CRC, 2002.

[10] Smith-Miles, Kate, and Leo Lopes. "Measuring instance difficulty for combinatorial optimization problems." Computers & Operations Research 39.5 (2012): 875-889.

[11] Christian Blum, Andrea Roli. “Metaheuristics in combinatorial optimization: Overview and conceptual comparison.” ACM Comput. Surv. 35, 3, pp. 268–308, 2003.

[12] I. V. Grebennik, O. S. Chorna, “Influence of Certain Transpositions on the Cyclic Structure of Permutations,” in Cybernetics and Systems Analysis, vol. 51(6), pp.947–955, 2015.

[13] I. V. Grebennik, O. S. Chorna, “Special Transpositions of Permutation Elements and Properties of Their Composition,” in Cybernetics and Systems Analysis, vol. 53(1), pp. 67–77, 2017.

[14] Lee, John A., and Michel Verleysen. Nonlinear dimensionality reduction. Vol. 1. New York: Springer, 2007.

[15] Yang, X.; Wang, Z.; Zhang, H.; Ma, N.; Yang, N.; Liu, H.; Zhang, H.; Yang, L. “A Review: Machine Learning for Combinatorial Optimization Problems in Energy Areas”. Algorithms, 15, 205, 2022.

[16] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. “Machine learning for combinatorial optimization: a methodological tour d’horizon”. arXiv:1811.06128, 2018.

[17] Maxime Gasse, Didier Chetelat, Nicola Ferroni, Laurent Charlin, Andrea Lodi “Exact Combinatorial Optimization with Graph Convolutional Neural Networks” Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada.

[18] Yu.G. Stoyan, S.V. Yakovlev, “Properties of convex functions on permutation polyhedron,” Reports of the Academy of Sciences of Ukraine, vol. 3 (A), pp.238–240, 1988.

[19] Stoyan, Y.G., Yakovlev, S.V. Theory and Methods of Euclidian Combinatorial Optimization: Current Status and Prospects. Cybern Syst Anal 56, 366–379 (2020).