# Low-Dimensional Euclidean Embedding for Visualization of Search Spaces in Combinatorial Optimization

Krzysztof Michalak
Department of Information Technologies, Institute of Business Informatics, Wroclaw University of Economics
Wroclaw, Poland
krzysztof.michalak@ue.wroc.pl,ORCiD:0000-0003-4994-9930

## ABSTRACT

This abstract summarizes the results reported in the paper [3]. In this paper a method named Low-Dimensional Euclidean Embedding (LDEE) is proposed, which can be used for visualizing high-dimensional combinatorial spaces, for example search spaces of metaheuristic algorithms solving combinatorial optimization problems. The LDEE method transforms solutions of the optimization problem from the search space $\Omega$ to $\mathbb{R}^k$ (where in practice $k = 2$ or 3). Points embedded in $\mathbb{R}^k$ can be used, for example, to visualize populations in an evolutionary algorithm.

The paper shows how the assumptions underlying the the t-Distributed Stochastic Neighbor Embedding (t-SNE) method can be generalized to combinatorial (for example permutation) spaces. The LDEE method combines the generalized t-SNE method with a new Vacuum Embedding method proposed in this paper to perform the mapping $\Omega \to \mathbb{R}^k$.

## CCS CONCEPTS

• **Computing methodologies → Scientific visualization**; • **Mathematics of computing → Combinatorial optimization**; *Evolutionary algorithms*;

## KEYWORDS

Visualization, Combinatorial optimization, Euclidean embedding, t-Distributed Stochastic Neighbor Embedding (t-SNE)

**ACM Reference Format:**
Krzysztof Michalak. 2019. Low-Dimensional Euclidean Embedding for Visualization of Search Spaces in Combinatorial Optimization. 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/10.1145/3319619.3326761

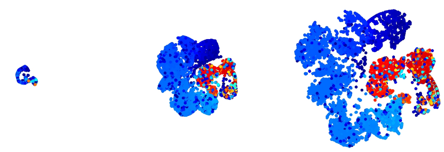

**Figure 1: The generalized t-SNE method working on solutions of the Four Peaks Problem (4PP) [1] for $n = 100$. Iterations: 200, 400, and 1000.**

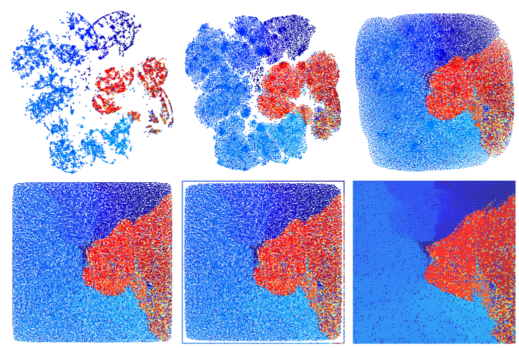

**Figure 2: The Vacuum Embedding method working on solutions of the Four Peaks Problem (4PP) for $n = 100$. In the top row: the embedding produced by t-SNE (left) and produced after 10 and 40 iterations of the Vacuum Embedding method. In the bottom row: iteration 500 of the Vacuum Embedding (the last one before edge "snapping" became active) and iterations in which edge "snapping" was active: 501, and the last one 582.**

## 1 METHOD DESCRIPTION

The Low-Dimensional Euclidean Embedding (LDEE) method consists of the t-Distributed Stochastic Neighbor Embedding (t-SNE) method [4] generalized to combinatorial (e.g. permutation) spaces and the Vacuum Embedding method. The method proposed in the paper is very general and can be applied to various combinatorial search spaces, most importantly binary vector spaces and permutation spaces.

**The working of the LDEE method**

The steps which constitute the LDEE method are as follows.

(1) An optimization algorithm generates solutions which are elements of the search space $\Omega$, for example permutations of a given length.
(2) The generalized t-SNE method is used to map solutions from $\Omega$ to $\mathbb{R}^k$, where, for visualization purposes, the dimensionality $k$ is 2 or 3 (Figure 1).
(3) The Vacuum Embedding (VE) method proposed in this paper places the points embedded in $\mathbb{R}^k$ on a regular square grid (Figure 2). It moves points by simulating forces exerted by


*GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic*

ACM ISBN 978-1-4503-6748-6/19/07.
https://doi.org/10.1145/3319619.3326761

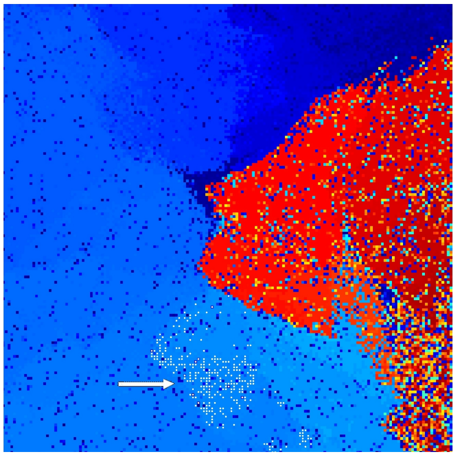

**Figure 3: Population of an evolutionary algorithm solving the Four Peaks Problem (white dots) at generation number 100. The population consists of good non-REWARDed solutions. Only after REWARDed solutions are found the population transits to the area with high values of evaluation function (coloured red).**

empty spaces in the grid and includes a "snapping" mechanism which aligns points to the grid starting from outside.

Overall, a mapping from $\Omega$ to $\mathbb{R}^k$ is performed which can be denoted as a composition of two transformations: $x'$ = LDEE($x$) = VE(t-SNE($x$)), where $x \in \Omega$ and $x' \in \mathbb{R}^k$. The steps 1-3 do not change attributes of solutions, such as the value of the objective function(s), values of constraint violation in constrained problems, information about genetic operators that generated the solutions, etc. Visualizations are made using positions of points in $\mathbb{R}^k$ corresponding to solutions in $\Omega$ obtained using the LDEE method. Attributes of solutions, such as objective function values can be used, for example, to colour the points plotted in the visualization. For example, on the regular grid produced by the LDEE method, populations of an evolutionary algorithm can be plotted (Figure 3).

**Main characteristics of the LDEE method**

Thanks to the properties of the t-SNE method, nearby points in $\Omega$ (which can be, for example, a permutation space with the Kendall-$\tau$ distance) are mapped to nearby points in $\mathbb{R}^k$. If it is desirable to preserve the distances as best as possible the next step may be omitted, but if the goal is to make every solution clearly visible the Vacuum Embedding method can be used to obtain a square grid. Some of the advantages of the presented approach are:

(1) A 1-1 mapping between solutions in $\Omega$ and points in $\mathbb{R}^k$ is preserved, allowing to plot every solution using e.g. a colour corresponding to the objective function value. It makes it easy to visually assess how many of the solutions have a certain characteristic, such as a low objective function value.
(2) The regular grid produced by the LDEE method is a good basis for plots such as the ones presented in this abstract, but also for example three-dimensional plots.
(3) Because nearby points in $\Omega$ are mapped to nearby points in $\mathbb{R}^k$ it is possible to assess how distances between solutions influence the values of the attributes of these solutions.

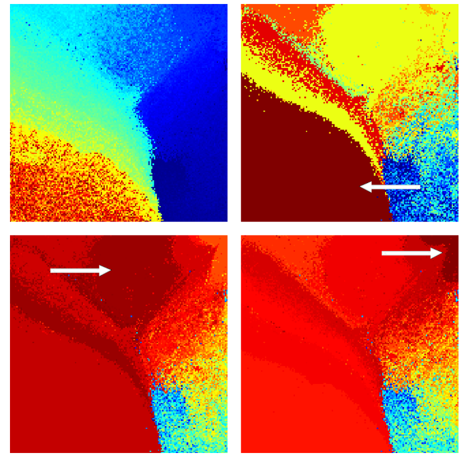

**Figure 4: Solutions of the bi-objective Firefighter Problem [2] with $N_v$ = 500 graph nodes. Plots show the generation number (top-left), the number of non-burning nodes (top-right) and the values of both objectives (bottom). Arrows show the areas with the highest values.**

(4) Contrary to methods that produce clusters of solutions or graph nodes the LDEE method makes all the solutions visible. This is particularly useful for visualizing population dynamics, because all solutions can be tracked and it is possible to see if they tend to group or scatter.

**The LDEE visualization for multiobjective optimization**

In Figure 4 solutions of the bi-objective Firefighter Problem [2] are presented. Clearly, the algorithm managed to increase the number of graph nodes which fire did not reach (top-right). Solutions maximizing individual objectives were found relatively early (two bottom plots). Because the algorithm performed Pareto optimization, it can be concluded that in later generations better trade-offs were found, not necessarily maximizing individual objectives.

## 2 ADDITIONAL MATERIALS

Additional materials can be found on the LDEE method web page at http://krzysztof-michalak.pl/research/ldee/index.html. This web page contains ready to use tools with which it is possible to apply the LDEE method to the user's own data set. A detailed example is available at http://krzysztof-michalak.pl/research/ldee/example.html along with data files and instructions on how to run the programs to obtain the presented results.

## REFERENCES

[1] Shumeet Baluja and Rich Caruana. 1995. Removing the Genetics from the Standard Genetic Algorithm. In *Machine Learning Proceedings 1995*, Armand Prieditis and Stuart Russell (Eds.). Morgan Kaufmann, San Francisco, 38–46.
[2] Krzysztof Michalak. 2015. The Sim-EA Algorithm with Operator Autoadaptation for the Multiobjective Firefighter Problem. In *Evolutionary Computation in Combinatorial Optimization*, Gabriela Ochoa and Francisco Chicano (Eds.). LNCS, Vol. 9026. Springer, 184–196.
[3] K. Michalak. 2019. Low-Dimensional Euclidean Embedding for Visualization of Search Spaces in Combinatorial Optimization. *IEEE Transactions on Evolutionary Computation* 23, 2 (2019), 232–246.
[4] L.J.P. van der Maaten. 2014. Accelerating t-SNE using Tree-Based Algorithms. *Journal of Machine Learning Research* 15 (2014), 3221–3245.