Edge Bundling as a Multi-Objective Optimization Problem

Raissa dos Santos Vieira
Universidade Federal de Goiás, Goiânia, Brazil

Hugo A. D. do Nascimento
Universidade Federal de Goiás, Goiânia, Brazil

Joelma de Moura Ferreira
IPOG Instituto de Pós-Graduação e Graduação, Goiânia, Brazil

Les Foulds
Universidade Federal de Goiás, Goiânia, Brazil

Karsten Klein
University of Konstanz, Germany

Falk Schreiber
University of Konstanz, Germany
Monash University, Melbourne, Australia

Abstract

Edge bundling is a technique commonly used to reduce visual clutter and improve the comprehension of the drawings of large graphs. Here, we model edge bundling as a multi-objective optimization problem and employ clustering strategies, metaheuristic and Pareto analysis to identify non-dominated solutions for some classical graphs from the literature.

2012 ACM Subject Classification Computing methodologies

Keywords and phrases Graph Drawing, Edge Bundling, Visual Clutter, Multi-objective Optimization

Digital Object Identifier 10.4230/LIPIcs.GD.2025.53

Category Poster Abstract

Funding H. A. D. do Nascimento, K. Klein, and F. Schreiber acknowledge funding by DFG under Germany's Excellence Strategy – EXC 2117 – 422037984, and DFG project ID 251654672 – TRR 161.

Acknowledgements We thank the computer support from LaMCAD/UFG, Brazil.

1 Introduction

A large variety of methods exist that realize edge bundling, see e. g. [3, 7, 4]. A formalization of some explicit edge bundling optimization problems has been proposed in [1]. Among those problems, the General-Based Edge Bundling (GBEB) Problem was introduced that aims at minimizing the number of bundles and at maximizing a given set of compatibility measures. The authors investigated in [1] a specific version of the GBEB, which consists of maximizing a weighted sum of two objective components: 1) the reciprocal of the number of bundles; and 2) the sum of the compatibility value for every bundle, calculated using the product of distinct compatibility measures for each pair of edges grouped together and a threshold to regulated the signal of the component. An evolutionary algorithm for the problem was designed and achieved good results. In a further work presented in [6], the authors treated the GBEB as a clustering problem and employed a combination of clustering and clustering-ensemble methods to solve it. The new approach led to better solutions than

33rd International Symposium on Graph Drawing and Network Visualization (GD 2025).
Editors: Vida Dujmović and Fabrizio Montecchiani; Article No. 53; pp. 53:1–53:5
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

when using the evolutionary algorithm. However, both studies adopted an objective function that combines distinct objectives in a fixed way, thus limiting the space of possibly desirable solutions.

This paper investigates the GBEB problem as a multi-objective optimization problem, by searching for the Pareto frontier considering all objectives independently, instead of combining them in a single objective function. The multi-objective version of GBEB (abbreviated here as MEB) is still treated as a clustering problem, as done in [6], but we go beyond the previous methodological approaches: a step-by-step method is adopted that combines bundling strategies with a meta-heuristic to extensively explore the solution space and find good compromise solutions.

2 The Multi-objective Approach

We aim at finding clustering solutions that maximize seven distinct aspects: 1) the reciprocal of the number of edge bundles (number of edge clusters), 2)-5) four compatibility measures proposed by Holten in [3] (angle, scale, position and visibility compatibilities), 6) the distance compatibility presented in [1], and 7) a new topological compatibility measure, which is defined as the distance between edge embeddings obtained with the *node2vec* method [2]. Note that only aspects 2) to 7) are compatibility measures. Here, the measurements in every quality aspect are normalized with maximum value =1.0.

The well known concept of *dominance* is used for comparing clustering solutions. A set of solutions can be divided into a hierarchical sequence of *dominance classes* by the following recursive procedure: solutions not dominated by any other solution in the set comprises the first class (called the *non-dominated class*); then, these solutions are removed from the set and the process repeats for the remaining set. While a class contains only equivalent solutions, any class is considered better than the subsequent classes. The first class constitutes the *Pareto frontier*. Our optimization method inputs a given graph $G = (V, E)$ and a straight-line drawing D of G, and performs the following three steps, where the first two are similar to what was done in [6]:

- **Step 1 – Creating edge representations**: Given G and D, an edge-to-edge compatibility matrix (dimension $|E| \times |E|$) is created for every compatibility measure. The i-th line of the j-th matrix, with $1 \leq i \leq |E|$ and $1 \leq j \leq 6$, is the *feature representation* of edge i according to the compatibility measure j.
- **Step 2 – Producing first solutions via clustering algorithms**: Four standard clustering algorithms (Agglomerative Hierarchical, K-Means, K-Means with Mini Batch, and Spectral Clustering) with various parameter configurations are employed to cluster the graph edges for each compatibility measure individually, using their related edge feature representation. A normalized quality vector (with seven values, representing the measurement for seven previously mentioned quality aspects) is computed for every solution. Finally, the 1000 solutions with the quality vector closest to $[1, 1, 1, 1, 1, 1, 1]$ are taken to the next step.
- **Step 3 – Improving solutions via multi-objective consensus**: This step runs an Asynchronous Team (A-Team) of autonomous agents [5] to improve the clustering solutions generated in Step 2, as an evolving group of dominance classes. The A-Team architecture consists of a shared memory with a predefined number N, of clustering solutions, and eighteen agents, which were formed by three clustering consensus algorithms with different setup parameters. Each agent reads up to k solutions randomly chosen from the memory (k is also predefined), runs a clustering consensus strategy, and writes

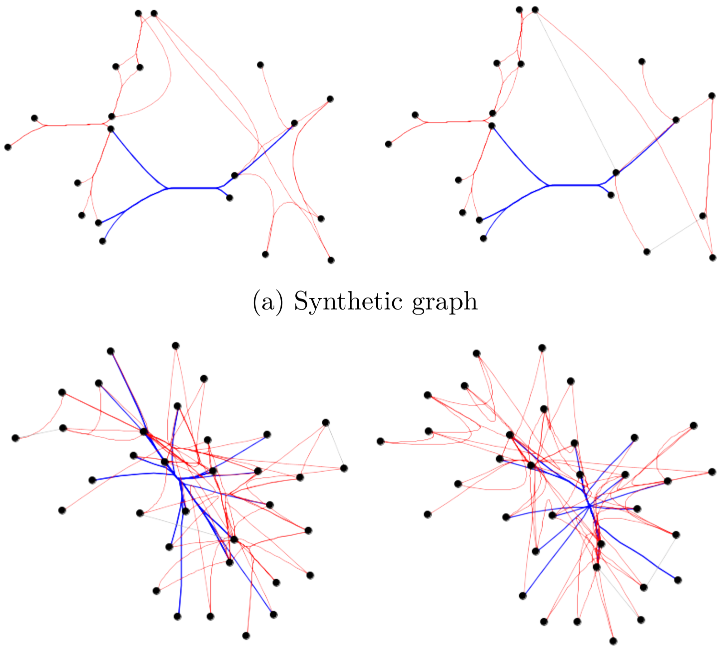
its resultant solution to the shared memory. Every new solution produced may cause the dominance classes to be recalculated and an earlier solution (preferably from the latest dominance classes) to be removed. The A-Team runs until a pre-specified stop condition is reached, when the non-dominated class of solutions (the final Pareto frontier) is outputted.

The solutions in the Pareto frontier can be graphically rendered by using the adapted force-directed edge bundling method described in [1]. It takes the graph, the position of its vertices and a list of sets of edges (clusters) and produces an edge-bundling graph drawing.

3 Experiments and Results

In order to validate the new multi-objective optimization approach, experiments were performed with five sparse graphs from the literature, with sizes ranging from 20 to 160 vertices, and from 28 to 161 edges. We also set $N = 1000$ and $k = 10$ in Step 3, and ran the A-Team algorithm for one hour (this was sufficient for producing good quality solutions).

As the main result, the A-Team algorithm was capable of improving the set of solutions by finding new non-dominated ones. It also reduced the size of the non-dominated class for all graphs, demonstrating that many solutions generated in Step 2 could be replaced by a few better ones. Another important contribution of the general approach is the diversification of compromise solutions, with regard to the distinct quality aspects, which becomes ready for human evaluation, as illustrated in Figure 1 (blue lines indicate the largest bundle).



(a) Synthetic graph

(b) Zachary Karate Club graph

Figure 1 Examples of compromise solutions in the Pareto frontier for two of the five graphs.

References

1 Joelma De M. Ferreira, Hugo A. D. Do Nascimento, and Les R. Foulds. An evolutionary algorithm for an optimization model of edge bundling. *Information*, 9(7), 2018. doi:10.3390/info9070154.

2 Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. *CoRR*, abs/1607.00653, 2016. doi:10.48550/arXiv.1607.00653.

3 Danny Holten and Jarke J. van Wijk. Force-directed edge bundling for graph visualization. *Comput. Graph. Forum*, 28(3):983–990, 2009. doi:10.1111/j.1467-8659.2009.01450.x.

4 Quan Hoang Nguyen, Seok-Hee Hong, and Peter Eades. TGI-EB: A new framework for edge bundling integrating topology, geometry and importance. In Marc J. van Kreveld and Bettina Speckmann, editors, *Graph Drawing - 19th International Symposium, GD 2011, Eindhoven, The Netherlands, September 21-23, 2011, Revised Selected Papers*, volume 7034 of *Lecture Notes in Computer Science*, pages 123–135. Springer, Springer, 2011. doi:10.1007/978-3-642-25878-7_13.

5 S Talukdar, L Baerentzen, A Gove, and P De Souza. Asynchronous teams: Cooperation schemes for autonomous agents. *Journal of heuristics*, 4(4):295–321, 1998. doi:10.1023/A:1009669824615.

6 Raissa S. Vieira, Hugo A. D. do Nascimento, Joelma de Moura Ferreira, and Les R. Foulds. Clustering ensemble-based edge bundling to improve the readability of graph drawings. In *26th International Conference Information Visualisation, IV 2022, Vienna, Austria, July 19-22, 2022*, pages 21–26. IEEE, 2022. doi:10.1109/IV56949.2022.00013.

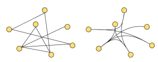
7 Markus Wallinger, Daniel Archambault, David Auber, Martin Nöllenburg, and Jaakko Peltonen. Faster edge-path bundling through graph spanners. *Comput. Graph. Forum*, 42(6):e14789, 2023. doi:10.1111/cgf.14789.

Edge Bundling as a Multi-objective Optimization Problem

Raissa dos Santos Vieira [1] Hugo A. D. do Nascimento [1] Joelma de Moura Ferreira [2]
Les Foulds [1] Karsten Klein [3] Falk Schreiber [3]
[1]Universidade Federal de Goiás, Brazil [2]IPOG Instituto de Pós-Graduação e Graduação, Brazil [3]Konstanz University, Germany

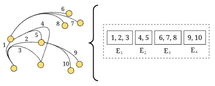
Introduction

Edge bundling is a technique commonly used to reduce visual clutter by grouping edges in bundles.



General-Based Edge-Bundling (GBEB)

Consists of an explicit edge bundling optimization problem that aims to minimize the number of bundles, and to maximize a set of compatibility measures. Solved using an evolutionary algorithm [1].



Clustering Ensemble-Based Edge Bundling Approach

Treats GBEB as a clustering problem, performing edge bundling through edge-clustering strategies followed by a clustering ensemble [2].

Max. $f = w_1 \times \frac{1}{n} + w_2 \times C_G$

with n = number of bundles, T a threshold, and $C_G = \sum_{i=1}^{n} C_{E_i}$

$$C_{E_i} = \begin{cases} \sum C(p,q), \forall p,q \in E_i, & \text{if } min(C(p,q)) \geq T; \\ p_e, & \text{if } min(C(p,q)) < T; \\ 0, & \text{if } |E_i| = 1. \end{cases}$$

$$P_e = \begin{cases} \sum C(p,q) \cdot (-3), & \text{if } \sum C(p,q) > 0; \\ -3, & \text{if } \sum C(p,q) = 0. \end{cases}$$

$$C(p,q) = C_a(p,q) \times C_s(p,q) \times C_p(p,q) \times C_v(p,q) \times C_d(p,q)$$

Multi-objective Edge Bundling (MEB)

MEB is a variant of GBEB defined as:

Let $G = (V,E)$ be a simple graph and D a straight-line drawing of G. Let $C_1(p,q), C_2(p,q), ..., C_m(p,q)$ be m compatibility measures defined for every pair of edges $p,q \in E$ based on D. The multi-objective explicit edge bundling problem in decentralized bundles (MEB) consists of determining a decomposition of E into disjoint subsets $E_1, E_2, ..., E_n$, with $E = \cup_{i=1}^{n} E_i$ and $E_i \cap E_j \neq \emptyset$ for $1 \leq i,j \leq |E|$ and $i \neq j$, that maximizes the objective vector $\overline{F} = [F_B, C_1, C_2, ..., C_m]$, where $F_B = \frac{1}{n}$ and $C_k = \sum_{i=1}^{n} C_{k,E_i}$ for $k = 1, 2, \ldots, m$, with C_{k,E_i} a version of C_{E_i} that uses only the k-th compatibility measure.

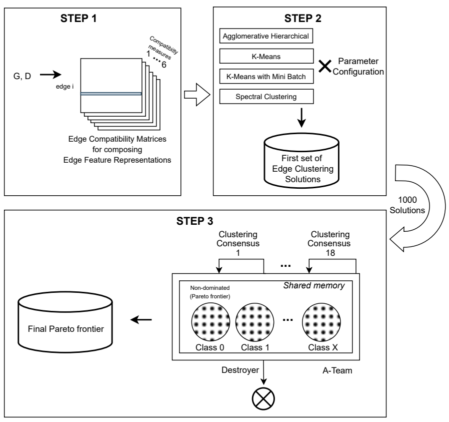
Additional explanation:

- The goal is to find a decomposition of E into disjoint subsets (clusters or bundles) that maximizes the values of $\overline{F}$.
- The number n of subsets of E is variable, with $1 \leq n \leq |E|$.
- Maximizing the vector $\overline{F}$ involves:
  1. Minimizing the number of bundles (n); and
  2. Maximizing m predefined compatibility measures.
- For the current study, $m = 6$ and the compatibility measures are *angle*, *scale*, *position*, *visibility*, *distance*, and *topological*.

Clustering Ensemble for Multi-objective Edge Bundling

To solve the multi-objective edge bundling problem, we present an approach that uses the edge-clustering strategies adopted in [2], but extend it to support Pareto Analysis and to include an Asynchronous Team (A-Team) algorithm.

The approach consists of the following steps:



Experiments and Results

Experiments were performed with five graphs, named G1 to G5. For the A-Team, we set $N = 1000$ as its memory size, and $k = 10$ as the maximum number of solutions for selection by any consensus agent. The A-Team ran for 3600 seconds for each graph.

Table 1 presents information about the graphs and the number of solutions in the Pareto frontier at Step 3. The evolution of the number of solutions on the Pareto frontier during the experiments can be observed in Figure 1. Figure 2 shows bundling solutions from the final Pareto frontier.

Graph	|V|	|E|	Pareto Frontier		No. of Iterations	New Solutions in Pareto Frontier
			Initial	*Final*		
G1 – Synthetic	20	28	325	35	22698	28
G2 – Zachary Karate Club	34	78	828	262	1908	16
G3 – Planar Graph GD2015	66	101	882	451	1260	20
G4 – Dolphin Social Network	62	159	958	903	198	65
G5 – MovieLens	160	161	758	688	198	16

Table 1. Graph attributes and the number of solutions in the Pareto frontier in Step 3 of the approach.

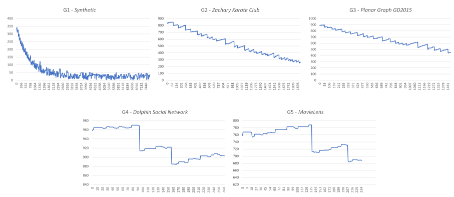
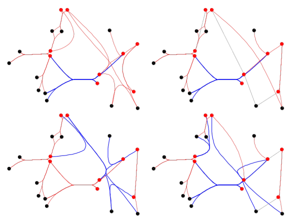
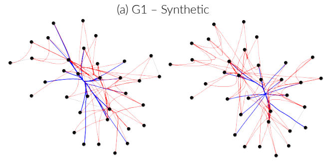


Figure 1. The change in the number of solutions in the Pareto frontier.



(a) G1 – Synthetic



(b) G2 – Zachary Karate Club

Figure 2. Examples of compromise solutions in the Pareto frontier for G1 and G2.

Main Results:

- A-Team was capable of improving the set of solutions by finding new non-dominated ones;
- The size of the non-dominated class for all graphs largely reduced over time, showing that many solutions generated in Step 2 could be replaced by better ones;
- The overall approach produces a set of compromise edge bundling solutions, with regard to the distinct quality aspects, thus supporting further human evaluation.

References

[1] Joelma De M. Ferreira, Hugo A. D. Do Nascimento, and Les R. Foulds. An evolutionary algorithm for an optimization model of edge bundling. *Information*, 9(7), 2018.

[2] Raissa dos Santos Vieira, Hugo A. D. do Nascimento, Joelma Moura Ferreira, and Les R. Foulds. Clustering ensemble-based edge bundling to improve the readability of graph drawings. In *Proceedings of 26th International Conference on Information Visualisation*, pages 21–26, 2022.