

Evaluating Visualizations of Sets and Networks that Use Euler Diagrams and Graphs

Almas Baimagambetov¹[0000-0002-5656-6324], Gem
Stapleton^{2,3}[0000-0002-6567-6752], Andrew Blake¹[0000-0001-5856-4544], and John
Howse¹[0000-0002-2329-2726]

¹ Centre for Secure, Intelligent and Usable Systems, University of Brighton, UK

{a.baimagambetov, john.howse, a.l.blake}@brighton.ac.uk

² University of Cambridge, Cambridge, UK, ges55@cam.ac.uk

³ University of Kent, Canterbury, UK, g.stapleton@kent.ac.uk

Abstract. This paper presents an empirical evaluation of state-of-the-art visualization techniques that combine Euler diagrams and graphs to visualize sets and networks. Focusing on SetNet, Bubble Sets and WebCola – techniques for which there is freely available software – our evaluation reveals that they can *inaccurately* and *ineffectively* visualize the data. Inaccuracies include placing vertices in incorrect zones, thus incorrectly conveying the sets in which the represented data items lie. Ineffective properties, which are known to hinder cognition, include drawing Euler diagrams with extra zones or graphs with large numbers of edge crossings. The results demonstrate the need for improved techniques that are more accurate and more effective for end users.

Keywords: Euler diagrams · Graphs · Sets · Networks · Visualization.

1 Introduction

A major goal of set and network data visualization is to draw diagrams that are both accurate and effective for users [1]. This is significant because there is a substantial amount of set and network data available, arising in various application areas [1, 18, 20], including bioinformatics, social networks, cartography and software architecture. This paper focuses on the common approach of using Euler diagrams and graphs in combination [1]. Euler diagrams represent sets and the graphs represent data items and relationships between them. We evaluate state-of-the-art techniques to reveal that they can produce layouts with both *inaccurate* and *ineffective* properties, hindering comprehension. No prior evaluation has compared set and network techniques by exposing the ways in which they possess inaccurate or ineffective properties. A take-away message from our research is that improved techniques are necessary for the effective visualization of sets and networks.

Section 2 summarises layout properties of Euler diagrams and graphs known to impact cognition, as well as existing techniques for drawing Euler diagrams

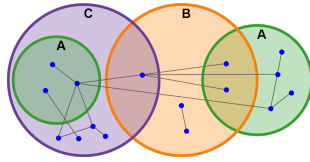


Fig. 1. SetNet.

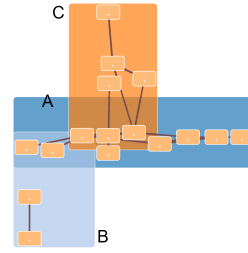


Fig. 2. WebCola.

and graphs in combination. A comparative evaluation of state-of-the-art techniques is given in section 3. We conclude and discuss future work in section 4.

2 Background

There are various layout properties of Euler diagrams and graphs that can lead to poor comprehension. Automated visualization techniques should aim to avoid such properties whilst ensuring an accurate visualization of the underlying data.

Euler Diagram Properties Prior research has led to the identification of properties of Euler diagrams that impact their effectiveness [3], with an important category being five *well-formedness properties* [15]:

- *Unique labels*: no two curves have the same label; curve labels that occur more than once are called *non-unique labels*. See Fig. 1, drawn using SetNet [16], where two curves labelled *A* represent the same set *A*.
- *Connected zones*: all of the zones in the diagram are connected components of the plane; a zone which is not connected is called *disconnected*.
- *Non-concurrent curves*: no parts of the curves run along the same path. In Fig. 2, drawn using WebCola [7], curves *A* and *B* share a concurrent segment.
- *Only two-points*: whenever a point is passed through by curves, it is passed through at most twice; points that fail this condition are called *triple points*.
- *Simple curves*: no curve self-intersects. Self-intersecting curves are *non-simple*.

Euler diagrams that break a well-formedness property hinder user comprehension [15] and possess one of the *ineffective properties*: non-unique labels, disconnected zones, concurrent curves, triple points, or non-simple curves. Other ineffective properties include *extra zones* and *non-circular curves*. Diagrams with extra zones are not *well-matched* to their semantics [9]. Circles were found to be a more effective curve shape than ellipses, squares or rectangles [3].

Diagrams are inaccurate when they omit zones that represent non-empty sets. Clearly, a diagram with *omitted zones* does not accurately reflect the data. In summary, automated Euler diagram layout techniques should avoid the ineffective properties of non-unique labels, disconnected zones, concurrent curves, triple points, non-simple curves, extra zones and non-circular curves, as well as the inaccurate property of omitted zones.

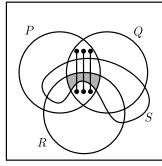


Fig. 3. Extra edge-curve crossings.

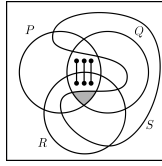


Fig. 4. No extra crossings.

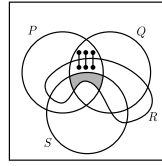


Fig. 5. No extra crossings.

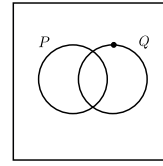


Fig. 6. A misplaced vertex.

Graph Properties Breaking the following properties reduces effectiveness [6]:

- *No edge crossings*: there are no points where two edges cross.
- *No edge bends*: edges are drawn as straight lines.
- *No edge-vertex intersections*: no edge passes through a non-incident vertex.
- *No vertex-vertex intersections*: there are no overlapping vertices.

That is, automated graph layout techniques should avoid the ineffective properties of edge crossings and edge bends and the inaccurate properties of edge-vertex intersections (where non-incident is required) and vertex-vertex intersections (which could appear as a single vertex).

Further Properties Combining Euler diagrams with graphs gives rise to further properties of interest:

- *No extra edge-curve crossings*: no graph edge passes through more curve segments than necessary.
- *Vertices in the correct zone*: no graph vertex lies on a curve or outside of the zone to which it belongs.

Graph edges connect vertices placed in Euler diagram zones. When vertices are in different zones, edges that connect them necessarily pass through some Euler diagram curves. Extra crossings between graph edges and the underlying diagram can lead to visual clutter, impairing readability [5]. In Fig. 3, the edges unnecessarily pass through the curve *S*; Figs. 4 and 5 redraw the diagram without the extra crossings. In addition, when vertices are located in the same zone, the Gestalt principle of *common region* [11] indicates that we perceive them as being grouped together, which should assist with a correct interpretation of the diagram. Extrapolating from this insight, the more curves passed through by an edge connecting two vertices may lead to a perception of them having less in common. In particular, the sets in which the data items, represented by the two vertices, both lie could be deemed fewer in number.

Properties and Layout Quality Summarising the above, visualization techniques should attempt to avoid properties that hinder cognition. Failure to avoid ineffective properties will give rise to diagrams that are ineffective. In addition, techniques should not produce diagrams that inaccurately represent the data from which they are derived. Such diagrams will lead to incorrect deductions being made about the underlying data.

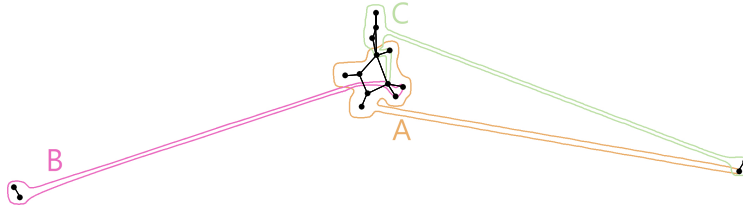


Fig. 7. Bubble Sets.

Existing Techniques A range of techniques draw combined Euler diagrams and graphs, including SetNet [16], WebCola [7], Bubble Sets [4], EulerView [19], Vizster [10] and KelpFusion [13]. A comprehensive overview of these and other techniques is available in [1]. Our evaluation, in section 3, focuses on SetNet, WebCola and Bubble Sets; justification for selecting these techniques is also provided in section 3.

SetNet [16] often produces diagrams with the ineffective property of non-unique labels, as in Fig. 1. *WebCola* [7] lays out the graph using a force-directed approach and then fits rectangles around the vertices. The rectangles attempt to have the smallest width and height needed to enclose their set members. This leads to diagrams with curve concurrency, as in Fig. 2 between *A* and *B*. *Bubble Sets* [4] routes curves around an already drawn graph. The graph is drawn first, independently of the curves, so Bubble Sets typically avoids ineffective graph properties. However, Fig. 7 allows us to observe that the graph layout can lead to convoluted curves. When routing the curves, Bubble Sets attempts to exclude non-set members within the curves, however this is not guaranteed. Hence, some visualizations are inaccurate. It is not possible for any technique to always accurately visualize the data and avoid all undesirable properties [8]. Our evaluation sets out to reveal the extent to which SetNet, Bubble Sets and WebCola produce diagrams with ineffective or inaccurate properties.

3 Evaluation: Inaccurate and Ineffective Properties

We selected techniques for our evaluation such that (a) they drew combined Euler diagrams and graphs, (b) the software was freely available, and (c) they could theoretically visualize any finite collection of sets and associated network, even though the implementation may fail. Bubble Sets [4], SetNet [16] and WebCola [7] were the techniques that met these criteria. All the diagrams used in our study are available at <https://github.com/AlmasB/D2020> and are marked to show where property violations occurred. The software implementations use different rendering details, such as vertex sizes, and different heuristics when producing algorithms from underlying theory. Note that such implementation decisions may impact the results presented in this paper. For example, changing the graph algorithm used for Bubble Sets, or the vertex sizes in WebCola, will likely affect the occurrences of inaccurate and ineffective properties.

3.1 Data for Visualization

In order to evaluate the techniques, we needed data for visualization. We used SNAP Twitter ego-networks [12] where data sets had up to 68413 vertices, 1685163 edges and 99 sets, which is too complex for the practical evaluation of visualization tools. The size of the data sets needed to be controlled and, as such, we appealed to existing evaluations to inform us of appropriate numbers of sets, vertices and edges. Consistent with previous studies [14, 15], we used a maximum of 8 sets. We selected the minimum number of sets to be two since, unlike in set-only data, a variety of combined Euler diagrams and graphs can be produced with two sets. Studies involving graphs typically had around 10 to 100 vertices and 40 to 170 edges [14, 16, 13, 17]. We selected data sets with 2, 4, 6, or 8 sets, 10 to 100 vertices, and 40 to 170 edges, to give controlled variety.

We then identified SNAP data with required numbers of sets and reduced the number of vertices and edges within them, to manage the network complexity. Firstly, we removed vertices with degree 0 as they do not materially affect any of the properties being evaluated. Secondly, we removed any multiple edges since, when drawn, the associated edges would just be on top of each other, and we also removed loops. Thirdly, we randomly removed vertices using an iterative approach, whilst ensuring that if this created any degree 0 vertices, they would be removed also until the number of vertices was in the given bounds. This process left us with a set of reduced SNAP data sets, from which we randomly selected a sample for our evaluation.

We explain how we selected reduced-complexity data sets for the 2-set case; the other cases are similar. We computed the median numbers of zones, vertices and edges. These medians were used to sub-divide the 2-set data sets into eight groups: a low (i.e. below the median) number of zones, vertices and edges; those with a low number of zones and vertices, but a high number of edges; those with a low number of zones and edges, but a high number of vertices; and so forth. From each of these eight sub-divisions, we randomly selected one data set for visualization, giving eight 2-set data sets. As we are visualizing 2-, 4-, 6-, and 8-set data sets, this gave us 32 data sets. In one case, SetNet could not draw the selected data set. This data set had 8 sets, a high number of zones, a low number of vertices and a high number of edges. No other SNAP data set could be reduced to match the given combination of the number of sets, zones, vertices and edges. The evaluation we present is based on the remaining 31 data sets.

3.2 Statistical Analysis Results

For each technique and for each data set, we counted the number of times each undesirable property occurred in the produced visualization. Our goal is to rank the techniques, for each undesirable property, to give an indication of relative effectiveness. For each technique, each undesirable property either (a) cannot be present, since the theoretical underpinnings of the technique ensured that it would not be, or (b) can be present. In category (a), the technique is necessarily superior, for that property, than any technique in category (b). For

Table 1. Means for inaccurate properties.

Property	SetNet	Bubble Sets	WebCola
Omitted zones	0	0	0.8
Edge-vertex intersections	76.2	6.9	213.0
Vertex-vertex intersections	2.2	0	2.5
Vertices in incorrect zone	0.9	0.5	2.6

Table 2. Means for ineffective properties.

Property	SetNet	Bubble Sets	WebCola
Non-unique labels	0.5	0	0
Disconnected zones	0	8.5	0.2
Concurrent curves	0	1.0	2.1
Triple points	0	0.7	0.5
Non-simple curves	0	0	0
Non-circular curves	0	4.9	4.9
Extra zones	0.1	1.9	0.5
Edge crossings	1236.8	588.4	1153.6
Extra edge-curve crossings	34.8	224.6	143.0

those in category (b), statistical analysis was performed on the counts in order to derive further ranking information. Given that the counts are not normally distributed and that the same data sets were used across four techniques, a non-parametric Friedman test was applied to rank at least three techniques and a Wilcoxon test was applied to rank exactly two techniques. If significant differences were revealed in the former case, a Nemenyi post-hoc analysis was used to derive a ranking. Throughout, results are taken to be significant at the 5% level. For all counts from which the statistical analysis results were obtained, see <https://github.com/AlmasB/D2020>. The techniques are abbreviated as follows: SetNet (SN), Bubble Sets (BS), and WebCola (WC).

Evaluation of Inaccurate Properties Table 1 summarises the mean counts of the inaccurate properties; the means for category (a) techniques are in bold and are necessarily 0. Table 3 shows which techniques participated in a statistical test, the p -value associated with the test and the derived post-hoc ranking; $A < B$ means A had significantly fewer inaccuracies, and therefore A is more

Table 3. The p -values and technique rankings for inaccurate properties.

Property	Statistical Test	p -value	Ranking
Omitted zones	–	N/A	BS = SN < WC
Edge-vert. inter.	SN, BS, WC	< 0.001	BS < SN < WC
Vert.-vert. inter.	SN, BS, WC	0.040	BS = SN = WC
Vert. incorrect zone	SN, BS, WC	0.035	BS < WC, BS = SN, SN = WC

Table 4. The p -values and technique rankings for ineffective properties.

Property	Statistical Test	p -value	Ranking
Non-unique labels	–	N/A	BS = WC < SN
Disconnected zones	BS, WC	0.001	SN < WC < BS
Concurrent curves	BS, WC	0.001	SN < BS < WC
Triple points	BS, WC	0.255	SN < BS = WC
Non-simple curves	–	N/A	SN = BS = WC
Non-circular curves	–	N/A	SN < BS = WC
Extra zones	SN, BS, WC	0.003	SN < BS, SN = WC, BS = WC
No edge crossings	SN, BS, WC	< 0.001	BS < SN = WC
No extra edge-curve crossings	SN, BS, WC	< 0.001	SN < BS = WC

accurate than *B*. Bubble Sets is consistently ranked as most, or jointly most, accurate. WebCola was ranked as least accurate or jointly least accurate, performing particularly badly for edge-vertex intersections with the mean of 213.0.

Evaluation of Ineffective Properties Table 2 summarises the mean counts for the ineffective properties and Table 4 shows the technique rankings; the techniques employed straight line edges, so there were never edge bends. The results here reveal there is no clear cut ‘least ineffective’ technique, as measured by these property counts. However, WebCola still fairs particularly poorly, being ranked worst or joint worst for seven out of the nine properties. SetNet is ranking best or joint best for seven out of the nine properties. Lastly, Bubble Sets is ranked best or joint best three times and worst or joint worst on five occasions. A further point of note is that non-unique labels and disconnected zones are particularly undesirable ineffective properties [15] and should be avoided where possible, even if that means other properties are exhibited. Bubble Sets was ranked (joint) best for non-unique labels but the worst for disconnected zones. By contrast, SetNet was ranked worst for non-unique labels and best for disconnected zones.

4 Conclusion and Future Work

A major problem when producing visualizations of data is finding an accurate and effective layout. A lot of attempts have been made to draw combined Euler diagrams and graphs but the state-of-the-art has not typically attempted to avoid properties that are empirically justified to be ineffective. Moreover, existing techniques can also produce inaccurate visualizations, thus not giving a true representation of data. There are three key take-away messages from our evaluation. Firstly, WebCola performed particularly badly: it is, compared to SetNet and Bubble Sets, particularly inaccurate and ineffective. Focusing on accuracy, Bubble Sets and SetNet were on par, except for edge-vertex intersections where Bubble Sets was superior. Regarding ineffective properties, SetNet is mostly superior to Bubble Sets. Our suggestion, based on these results, is that Bubble Sets should be the technique of choice whenever visualization accuracy is more important than effectiveness.

An important factor to consider is that the evaluated properties act as a proxy to diagram effectiveness. We should be mindful that avoiding inaccurate and ineffective properties does not necessarily ensure effectiveness. Our evaluation has focused on countable properties, yet there exist other properties, which relate to aesthetics, that can also impact effectiveness [2]. As it currently stands, the aesthetics of diagrams are not readily measurable and, so, were not part of our evaluation. In the future, empirical studies involving human participants are needed.

Acknowledgement Gem Stapleton was partially funded by a Leverhulme Trust Research Project Grant (RPG-2016-082) for the project entitled Accessible Reasoning with Diagrams and EPSRC grant EP/T019603/1.

References

1. Alsallakh, B., Micalef, L., Aigner, W., Hauser, H., Miksch, S., Rodgers, P.: The state-of-the-art of set visualization. *Computer Graphics Forum* **35**(1), 234–260 (2016)
2. Benoy, F., Rodgers, P.: Evaluating the comprehension of Euler diagrams. In: 11th Int. Conf. on Information Visualization. pp. 771–778. IEEE (2007)
3. Blake, A., Stapleton, G., Rodgers, P., Howse, J.: The impact of topological and graphical choices on the perception of Euler diagrams. *Information Sciences* **330**, 455 – 482 (2016).
4. Collins, C., Penn, G., Carpendale, M.S.T.: Bubble sets: Revealing set relations with isocontours over existing visualizations. *IEEE Trans. on Visualization and Computer Graphics* **15**(6), 1009–1016 (2009)
5. Debiasi, A.: Study of Visual Clutter in Geographic Node-Link Diagrams. Ph.D. thesis, University of Trento (2016)
6. Dunne, C., Shneiderman, B.: Improving graph drawing readability by incorporating readability metrics: A software tool for network analysts. Uni. of Maryland, HCIL Tech Report HCIL-2009-13 (2009)
7. Dwyer, T.: Webcola: Constraint-based layout in the browser. <https://ialab.it.monash.edu/webcola/> (2013). Last accessed 2 May 2019
8. Flower, J., Howse, J.: Generating Euler diagrams. In: *Diagrams*. pp. 61–75. Springer, (2002)
9. Gurr, C.: Effective diagrammatic communication: Syntactic, semantic and pragmatic issues. *J. of Visual Languages and Computing* **10**(4), 317–342 (1999)
10. Heer, J., Boyd, D.: Vizster: Visualizing online social networks. In: *IEEE Symp. on Information Visualization*. pp. 32–39. IEEE (2005)
11. Koffka, K.: *Principles of Gestalt Psychology*. Lund Humphries (1935)
12. Leskovec, J., Krevl, A.: SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data> (2014). Last accessed 9 Oct 2019
13. Meulemans, W., Henry Riche, N., Speckmann, B., Alper, B., Dwyer, T.: Kelpfusion: A hybrid set visualization technique. *IEEE Trans. on Visualization and Computer Graphics* **19**(11), 1846–1858 (2013)
14. Riche, N., Dwyer, T.: Untangling Euler diagrams. *IEEE Trans. on Visualization and Computer Graphics* **16**(6), 1090–1099 (2010)
15. Rodgers, P., Zhang, L., Purchase, H.: Wellformedness properties in Euler diagrams: Which should be used? *IEEE Trans. on Visualization and Computer Graphics* **18**(7), 1089–1100 (2012)
16. Rodgers, P., Stapleton, G., Alsallakh, B., Micalef, L., Baker, R., Thompson, S.: A task-based evaluation of combined set and network visualization. *Information Sciences* **367-368**, 58 – 79 (2016).
17. Saket, B., Simonetto, P., Kobourov, G., Borner, K.: Node, node-link, and node-link-group diagrams: An evaluation. *IEEE Trans. on Visualization and Computer Graphics* **12**(20), 2231–2240 (2014)
18. Santamaría, R., Therón, R., Quintales, L.: Bicoverlapper: a tool for bicluster visualization. *Bioinformatics* **24**(9), 1212–1213 (2008)
19. Simonetto, P., Auber, D., Archambault, D.: Fully automatic visualisation of overlapping sets. *Computer Graphics Forum* **28**(3), 967–974 (2009)
20. Wasserman, S., Faust, K.: *Social Network Analysis*. CUP (1994)