Report from Dagstuhl Seminar 19192

Visual Analytics for Sets over Time and Space

Edited by

Sara Irina Fabrikant[1], Silvia Miksch[2], and Alexander Wolff[3]

1 **Universität Zürich, CH,** sara.fabrikant@geo.uzh.ch
2 **TU Wien, AT,** miksch@ifs.tuwien.ac.at
3 **Universität Würzburg, DE, https://orcid.org/0000-0001-5872-718X**

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 19192 "Visual Analytics for Sets over Time and Space", which brought together 29 researchers working on visualization (i) from a theoretical point of view (graph drawing, computational geometry, and cognition), (ii) from a temporal point of view (visual analytics and information visualization over time, HCI), and (iii) from a space-time point of view (cartography, GIScience). The goal of the seminar was to identify specific theoretical and practical problems that need to be solved in order to create dynamic and interactive set visualizations that take into account time and space, and to begin working on these problems.

The first 1.5 days were reserved for overview presentations from representatives of the different communities, for presenting open problems, and for forming interdisciplinary working groups that would focus on some of the identified open problems as a group. There were three survey talks, ten short talks, and one panel with three contributors. The remaining three days consisted of open mic sessions, working-group meetings, and progress reports. Five working groups were formed that investigated several of the open research questions. Abstracts of the talks and a report from each working group are included in this report.

Seminar May 5–10, 2019 – http://www.dagstuhl.de/19192

2012 ACM Subject Classification Human-centered computing → Visualization, Theory of computation → Design and analysis of algorithms

Keywords and phrases Geovisualization, graph drawing, information visualization, set visualization, visual analytics

Digital Object Identifier 10.4230/DagRep.9.5.31

Edited in cooperation with Philipp Kindermann

1 Executive Summary

Sara Irina Fabrikant
Silvia Miksch
Alexander Wolff

Seminar Goals

Increasing amounts of data offer great opportunities to promote technological progress and business success. Visual analytics aims at enabling the exploration and the understanding of large and complex data sets by intertwining interactive visualization, data analysis, human-computer interaction, as well as cognitive and perceptual science. Cartography has for thousands of years dealt with the depiction of spatial data, and more recently geovisual analytics researchers have joined forces with the visual analytics community to create visualizations to help people to make better and faster decisions about complex problems that require the analysis of big data.

Set systems comprise a generic data model for families of sets. A *set* is defined as a collection of unique objects, called the set elements, with attributes, membership functions, and rules. Such a complex data model asks for appropriate exploration methods. As with many types of data, set systems can vary over time and space. It is important, however, not to treat time and space as usual variables. Their special characteristics such as different granularities, time primitives (time points vs. intervals), hierarchies of geographic or administrative regions need to be taken into account. Visualizing and analyzing such changes is challenging due to the size and complexity of the data sets.

Sets systems can also be seen as hypergraphs where the vertices represent the ground elements and the edges are the sets. However, compared to conventional graphs that represent only binary relations (that is, sets with two elements), the visualization of general hypergraphs has received little attention. This is even more so when dealing with dynamic hypergraphs or hypergraphs that represent spatial information.

In this seminar, we aimed at bringing together researchers from the areas of visual analytics, information visualization and graph drawing, geography and GIScience, as well as cartography and (spatial) cognition, in order to develop a theory and visualization methods for set systems that vary over time and space.

Seminar Program

As the topic of the seminar was interdisciplinary and the participants had very different scientific backgrounds, we introduced the main themes of the seminar in three separate sections: "Sets in Time", "Sets in Space", and "Graph Drawing and Set Visualization". Each section consisted of a survey talk and three to four short talks. The three sections were followed by a panel discussion. For the survey talks, we explicitly asked the presenters to give a balanced overview over their area (rather than to focus on their own scientific contributions).

On the second day of the seminar, we collected a number of challenging open problems. Then we formed five groups, each of which worked on a specific open problem for the remainder of the seminar. The work within the groups was interrupted only a few times; in order to share progress reports, listen to open-mic talks, and to discuss possible future activities. These plenary meetings helped to exchange the different visions of the working groups.

We now list the items of the program in detail.

1. Section "Sets in Time" (for abstracts, see Section 3)
   - Peter Rogers gave an excellent survey talk about techniques for visualizing sets over time. He illustrated possible challenges and opportunities in this research area.
   - Philipp Kindermann presented some results and open questions in simultaneous orthogonal graph drawing.
   - Wouter Meuleman talked about spatially and temporally coherent visual summaries.
   - Tamara Mchedlidze introduced a data-driven approach to quality metrics of graph visualizations.
   - Margit Pohl discussed perception considerations of space and time in cognitive psychology and their implications for the design of visualizations.

2. Section “Sets in Space” (for abstracts, see Section 4)
   - Sara Fabrikant gave an inspiring survey talk about space discussed from a cartographer‘s view.
   - Natalia Andrienko elaborated about evolving sets in space.
   - Somayeh Dodge discussed dynamic visualization of interaction in movement of sets.
   - Jan-Henrik Haunert introduced fast retrieval of abstracted representations for sets of points within user-specified temporal ranges.
3. Section “Graph Drawing and Set Visualization” (for abstracts, see Section 5)
   - André Schulz very nicely surveyed the area of drawing graphs and hypergraphs and sketched the main challenges in this area.
   - Michalis Bekos gave a short overview of graph drawing beyond planarity.
   - Sabine Cornelsen talked about general support for hypergraphs.
   - Martin Nöllenburg introduced plane supports for spatial hypergraphs.
4. The panel discussion was entitled “Visual Analytics for Sets over Time and Space: What are the burning scientific questions? An interdisciplinary perspective.” André Skupin, Steven Kobourov, and Susanne Bleisch each gave a short statement about the central questions of his or her area; see Section 6. Afterwards we had a fruitful and interesting discussion, which led to a productive open problem session.
5. The working groups formed around the following open problems:
   - “Concentric Set Schematization”,
   - “From Linear Diagrams to Interval Graphs”,
   - “Thread Visualization”,
   - “Clustering Colored Points in the Plane”, and
   - “Flexible Visualization of Sets over Time and Space”.

The reports of the working groups were collected by Michalis Bekos, Steven Chaplick, William Evans, Jan-Henrik Haunert, and Christian Tominski; see Section 7.

Future Plans

During our seminar, plans for a follow-up seminar were discussed in a plenary meeting. The seminar-to-be will aim at integrating the approaches for set visualization that have been taken by the different communities (geovisualization, information visualization, and graph drawing, including industry and research). Susanne Bleisch, Steven Chaplick, Jan-Henrik Haunert, and Eva Mayr are currently discussing the precise focus and a title to match that focus.

Among the 29 participants of the seminar, 24 participated in the survey that Dagstuhl does at the end of every seminar. Many answers were in line with the average reactions that Dagstuhl collected over a period of 60 days before our seminar (such as the scientific quality of the seminar, which received a median of 10 out of 11 – “outstanding”). A few questions, however, received different feedback. For example, due to the interdisciplinary nature of the seminar, we had more frequent Dagstuhl visitors than usually: a third of the participants of the survey had been to Dagstuhl at least seven times. It was also interesting to see that more participants than usually stated that our seminar had inspired new research ideas, joint projects or publications, that it had led to insights from neighboring fields, and that it had identified new research directions.

In spite of the organizers’ attempt to have a diverse group of participants, all survey participants were from academia and only two rated themselves as “junior”. Not surprisingly, some participants suggested to have more PhD students, more people from industry, and generally more people from applications rather than from (graph drawing) theory. The last free text comment in the survey reads: “Once again, a great week at Schloss Dagstuhl – thank you!”

Acknowledgments

We all enjoyed the unique Dagstuhl atmosphere. In particular, it was great to have the opportunity to use a separate room for each working group. We thank Philipp Kindermann for collecting the self-introductory slides before the seminar and for assembling this report after the seminar.

2 Table of Contents

3 Overview of Talks about "Sets in Time"

3.1 Techniques for Visualizing Sets over Time

Peter Rodgers (University of Kent – Canterbury, GB)

This talk surveyed the current work and potential new avenues when visualizing both set and time data simultaneously. The motivation for this work comes from set based data that changes over time in research areas such as Social Media, Biosciences and Medicine. We consider mental map preservation over effective layout, adding dynamic aspects to current set visualization methods, and scalability issues.

The initial sections of the talk looked at the state-of-the-art in set visualization and time visualization. The set visualization summary largely came from a survey [2] and briefly outlined the wide variety of set visualization methods, from Euler-like, region oriented, line based, glyph and node-link based, to name just a few. The overview of time visualization methods were, again, largely based on a survey [1]. Techniques for time visualization are less diverse than set visualization, being broadly classified into linear and cyclical methods. The survey then overviewed the few existing visualization techniques that can claim to visualize both time and sets: TimeSets [6], Time-Sets [5], Hypenet [8], Bubble Sets [3], Dynamic Euler Diagrams [7], Linear Representations [9], and Circos [4].

An important take home message is that the number of visualization methods that consider both sets and time is small. Hence, given the demand for such techniques this area is a potentially fruitful research area. We consider that developing dynamic versions of existing set visualizations would be a rich seam of new ideas. Merging current set and time visualizations is another promising route to visualizing this complex data.

References

1 W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Springer London, 2011. doi:10.1007/978-0-85729-079-3.

2 B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. The state-of-the-art of set visualization. *Computer Graphics Forum*, 35(1):234–260, nov 2015. doi:10.1111/cgf.12722.

3 C. Collins, G. Penn, and S. Carpendale. Bubble sets: Revealing set relations with iso-contours over existing visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009–1016, nov 2009. doi:10.1109/tvcg.2009.122.

4 M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra. Circos: An information aesthetic for comparative genomics. *Genome Research*, 19(9):1639–1645, jun 2009. doi:10.1101/gr.092759.109.

5 M. Masoodian and L. Koivunen. Temporal visualization of sets and their relationships using time-sets. In *2018 22nd International Conference Information Visualisation (IV)*. IEEE, jul 2018. doi:10.1109/iv.2018.00025.

6 P. H. Nguyen, K. Xu, R. Walker, and B. W. Wong. TimeSets: Timeline visualization with set relations. *Information Visualization*, 15(3):253–269, oct 2015. doi:10.1177/1473871615605347.

7 P. Rodgers, P. Mutton, and J. Flower. Dynamic euler diagram drawing. In *2004 IEEE Symposium on Visual Languages – Human Centric Computing*. IEEE, 2004. doi:10.1109/vlhcc.2004.21.

8 P. Valdivia, P. Buono, and J.-D. Fekete. Hypenet: Visualizing dynamic hypergraphs. In A. P. Puig and T. Isenberg, editors, *EuroVis 2017 – Posters*. The Eurographics Association, 2017. doi:10.2312/eurp.20171162.

9 T. von Landesberger, S. Bremm, N. Andrienko, G. Andrienko, and M. Tekusova. Visual analytics methods for categoric spatio-temporal data. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, oct 2012. doi:10.1109/vast.2012.6400553.

3.2 Simultaneous Orthogonal Graph Drawing

Philipp Kindermann (Universität Würzburg, DE)

Joint work of Patrizio Angelini, Steven Chaplick, Sabine Cornelsen, Giordano Da Lozzo, Giuseppe Di Battista, Peter Eades, Philipp Kindermann, Jan Kratochvíl, Fabian Lipp, Ignaz Rutter

Main reference Patrizio Angelini, Steven Chaplick, Sabine Cornelsen, Giordano Da Lozzo, Giuseppe Di Battista, Peter Eades, Philipp Kindermann, Jan Kratochvíl, Fabian Lipp, Ignaz Rutter: "Simultaneous Orthogonal Planarity", in Proc. of the Graph Drawing and Network Visualization – 24th International Symposium, GD 2016, Athens, Greece, September 19-21, 2016, Revised Selected Papers, Lecture Notes in Computer Science, Vol. 9801, pp. 532–545, Springer, 2016.

URL https://doi.org/10.1007/978-3-319-50106-2_41

We introduce and study the OrthoSEFE-k problem: Given k planar graphs each with maximum degree 4 and the same vertex set, is there an assignment of the vertices to grid points and of the edges to paths on the grid such that the same edges in distinct graphs are assigned the same path and such that the assignment induces a planar orthogonal drawing of each of the k graphs?

We show that the problem is NP-complete for $k \geq 3$ even if the shared graph is a Hamiltonian cycle and has sunflower intersection and for $k \geq 2$ even if the shared graph consists of a cycle and of isolated vertices. Whereas the problem is polynomial-time solvable for $k = 2$ when the union graph has maximum degree five and the shared graph is biconnected. Further, when the shared graph is biconnected and has sunflower intersection, we show that every positive instance has an OrthoSEFE-k with at most three bends per edge.

3.3 Spatially and Temporally Coherent Visual Summaries

Wouter Meulemans (TU Eindhoven, NL)

Joint work of Juri Buchmüller, Wouter Meulemans, Bettina Speckmann, Kevin Verbeek, Jules Wulms

When exploring large time-varying data sets, visual summaries are a useful tool to identify time intervals of interest for further consideration. A typical approach is to represent the data elements at each time step in a compact one-dimensional form or via a one-dimensional ordering. Such 1D representations can then be placed in temporal order along a time line. There are two main criteria to assess the quality of the resulting visual summary: (1) how well does the 1D representation capture the structure of the data at each time step, and (2) how coherent or stable are the 1D representations over consecutive time steps or temporal ranges? We focus on techniques that create such visual summaries using 1D orderings for time-varying spatial data. Specifically, we consider the case of moving 2D point objects.

We first analyze three *orientation-based* shape descriptors on a set of continuously moving points: the first principal component, the smallest oriented bounding box and the thinnest strip. If we bound the speed with which the orientation of the descriptor may change, this may lower the quality of the resulting shape descriptor. We first show that there is no *stateless algorithm*, an algorithm that keeps no state over time, that both approximates the minimum cost of a shape descriptor and achieves continuous motion for the shape descriptor. On the other hand, if we can use the previous state of the shape descriptor to compute the new state, then we can define "chasing" algorithms that attempt to follow the optimal orientation with bounded speed. Under mild conditions, we show that chasing algorithms with sufficient bounded speed approximate the optimal cost at all times for oriented bounding boxes and strips.

To compute visual summaries, we introduce a stable and efficient ordering for moving points which is based on principal components. Our method allows us to make an explicit trade-off between the two quality criteria. We conduct computational experiments that compare our method to various state-of-the-art approaches for computing 1D orderings for spatial data, based on a set of well-established quality metrics that capture the two main criteria. The experiments show that our *Stable Principal Component* (SPC) algorithm outperforms existing methods: the spatial quality of SPC is essentially equivalent to the methods that perform best for this criterion, the run time of SPC is a fast as the fastest methods, and SPC is more stable than all other methods tested.

3.4 A Data-Driven Approach to Quality Metrics of Graph Visualizations

Tamara Mchedlidze (KIT – Karlsruher Institut für Technologie, DE)

Joint work of Moritz Klammler, Tamara Mchedlidze, Alexey Pak
Main reference Moritz Klammler, Tamara Mchedlidze, Alexey Pak: "Aesthetic Discrimination of Graph Layouts", in Proc. of the Graph Drawing and Network Visualization – 26th International Symposium, GD 2018, Barcelona, Spain, September 26-28, 2018, Proceedings, Lecture Notes in Computer Science, Vol. 11282, pp. 169–184, Springer, 2018.
URL https://doi.org/10.1007/978-3-030-04414-5_12

Graph Visualization is a research area concerning automatic creation of pictorial representations of graphs. A node-link diagram (also called graph layout) is one of the most intuitive of these representations: the nodes are represented as 2 or 3-dimensional objects and edges as (poly-) lines or curves connecting the adjacent nodes. Node-link diagrams are used in a number of fields, including social science, bioinformatics, neuroscience, electronics, software engineering, business informatics and humanities.

Central to the Graph Visualization is the notion of the quality metric – a measure that formalizes how readable, clear and aesthetically pleasing a graph layout is. Some examples of simple quality metrics include number of edge crossings, edge crossing angle, drawing resolution. More complex quality metrics are the energy of a corresponding system of physical bodies and a linear combination of simple quality metrics. Quality metrics are utilized by network visualization algorithms to produce readable and aesthetically pleasing graph layouts.

In this talk I consider an alternative perspective on the quality metrics of graph layouts, by addressing the following question: "Of two given layouts of the same graph, which one is more aesthetically pleasing?". With that, I admit that "the ultimate" quality metric may not exist and one can hope for at most a (partial) ordering of layouts with respect to their aesthetic value. I introduce a neural network-based discriminator model trained on a labeled data set that decides which of two layouts has a higher aesthetic quality. The model demonstrates a mean prediction accuracy of 97.58%, outperforming discriminators based on an energy function and on the linear combination of popular quality metrics.

3.5 Perception of Space and Time – Implications for the Design of Visualizations

Margit Pohl (TU Wien, AT)

There is a considerable amount of research in cognitive psychology concerning the perception of space and time. Some of this research is relevant for the design of visualizations representing spatial and temporal data, although it should be mentioned that the application of basic research from psychology in visualization design is not always straightforward.

Human visual perception has several characteristics that are important for the design of visualizations. Gibson's ecological approach of visual perception implies that visual perception is related to movement in space. Perception is not a sequence of static pictures but a continuous flow of images while people move in the environment (optic flow field). When people are interested in objects in their environment they will move closer or go up and down a larger object (e.g., a house). In interfaces, such processes (zooming, panning, scrolling, ...) are mimicked. In contrast to other interaction possibilities, these are more natural and intuitive. This does not imply that less natural interaction possibilities are not effective, but they are less intuitive and have to be learned.

When people navigate in an environment they tend to develop schematic mental models of their environments. These mental models are incomplete and might be erroneous. Nevertheless, human navigation is generally very successful because they use information from the environment to continuously adapt their navigation (situatedness of spatial mental models).

Results from research on the perception of time is less relevant for the design of visual representations of temporal information. One important aspect in this context is the fact that humans use space as a metaphor for the representation of time. Using timelines is a very common way to reason about time. Another possibility is the usage of animation, although this should be designed appropriately. A specific challenge in the context of the visualization of spatial and temporal information is the combination of those two. When space is used as a metaphor for temporal information, this might not be compatible with the representation of geographic information. Especially in the context of very complex data, the usage of animation might be advisable.

4 Overview of Talks about "Sets in Space"

4.1 Evolving Sets in Space

Natalia V. Andrienko (Fraunhofer IAIS – Sankt Augustin, DE) and Gennady Andrienko (Fraunhofer IAIS – Sankt Augustin, DE)

Topic 1: Evolving spatial clusters (sets) of point events. Sets emerge, grow and shrink (cardinality and/or spatial extent), merge, split, disappear. Currently we use 3 visual displays: animated map, space-time cube, and bars along a time line (Gantt chart); each shows only a part of the information. Problem: the views are hard to link for getting the full picture. How to support visual exploration in a better way?

Topic 2: Subgroups in coordinated movement of multiple entities (e.g. shoal of fishes, football players). Synchronous coordination: continuous changes; subgroups of similarly moving entities emerge and change over time (e.g., entities separate from a group and join other groups). Problem: how to support visual exploration of subgroup formation and evolution?

Asynchronous coordination: multiple actors perform a sequence of activities towards a common goal, e.g., football players perform a sequence of passes during a game. Groups exist over certain time intervals. Changes are discrete: from interval to interval. Problem: how to support detecting repeated patterns of grouping and understanding the contexts in which they occur.

4.2 Dynamic Visualization of Interaction in Movement of Sets

Somayeh Dodge (University of Minnesota – Minneapolis, US)

Movement is a spatiotemporal process which involves space, time, and context. This presentation highlighted methods to visualize movement of sets and their interaction in space and time. 2D and 3D dynamic and interactive visualizations are created to highlight interaction among moving entities using an open-source visualization package, called DYNAMOvis to capture patterns of interaction. Direct interaction of entities are identified as the proximity of entities in space and time using spatial and temporal buffers. Visual variable color is used to highlight when entities are close together or meet at the same location and time. The geographic of context of movement and the interaction is visualized as background maps and satellite images. Time is captured using animation and the third dimension of space-time cube. As a a case study the presentation showed how the methods can be applied to highlight interaction between two tigers with adjacent home ranges.

4.3 Interactive Exploration of Spatio-Temporal Point Sets

Jan-Henrik Haunert (Universität Bonn, DE)

Joint work of Annika Bonerath, Jan-Henrik Haunert, and Benjamin Niedermann
Main reference Annika Bonerath, Jan-Henrik Haunert, and Benjamin Niedermann: "Computing alpha-shapes for temporal range queries on point sets". EuroCG 2019.
URL http://www.eurocg2019.uu.nl/papers/28.pdf

In my talk I present an overview of current developments of the Geoinformation Group at the University of Bonn, including algorithmic approaches to map generalization and label placement. A focus of my talk will be a new approach for the interactive exploration of spatio-temporal data. Generally, the aim is to develop data structures that can be repeatedly queried to obtain simple visualizations of parts of the data. In particular, the data is assumed to be a set of points each associated with a time stamp and the result of each query is visualized by an α-shape, which generalizes the concept of convex hulls. Instead of computing each shape independently, a simple data structure is proposed that aggregates the α-shapes of all possible queries. Once the data structure is built, it particularly allows a user to query single α-shapes without retrieving the actual (possibly large) point set.

5 Overview of Talks about "Graph Drawing and Set Visualization"

5.1 Survey on Graph and Hypergraph Drawing

André Schulz (FernUniversität in Hagen, DE)

Over the last 30 years graph drawing became an active area of research that bridges problems from graph theory and computational geometry. I present some classical graph drawing problems and discuss quality measures, prominent graph classes and drawing styles. Then a few selected representative graph drawing problems and algorithms are explained.

The second part of the talk covers hypergraphs and their connections to set visualization. I explain the differences between subset-based, partition-based and edge-based methods. I highlight the tight border between feasible and infeasible problems on the example on displaying 2 partitions simultaneously.

5.2 A Short Overview of Graph Drawing Beyond Planarity

Michael Bekos (Universität Tübingen, DE)

Beyond planarity is a new research direction in Graph Drawing, which is currently receiving increasing attention. Its primary motivation stems from recent cognitive experiments showing that the absence of specific kinds of edge-crossing configurations has a positive impact on the human understanding of a graph drawing. Graph drawing beyond planarity is concerned with the study of non-planar graphs that can be drawn by locally avoiding specific edge-crossing configurations or by guaranteeing specific properties for the edge crossings.

In this context, several classes of "beyond-planar graphs" have been introduced and studied, e.g., k-planar graphs, k-quasi planar graphs, right-angle-crossing graphs, fan-planar graphs, and fan-crossing free graphs. These classes of graphs have been mainly studied both in terms of their combinatorial properties and in terms of algorithms able to recognize and draw them.

In this talk, we will give a short overview of this new research direction aiming at covering a range of topics that are concerned with aspects of graph drawing and network visualization beyond planarity, such as combinatorial aspects of beyond-planar graphs, relationships between classes of beyond-planar graphs, and the complexity of the recognition problem for certain classes of beyond-planar graphs.

5.3 Supports for Hypergraphs

Sabine Cornelsen (Universität Konstanz, DE)

Joint work of Jan Christoph Athenstädt, Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, Arnaud Sallaberry
Main reference Jan Christoph Athenstädt, Sabine Cornelsen: "Planarity of Overlapping Clusterings Including Unions of Two Partitions", J. Graph Algorithms Appl., Vol. 21(6), pp. 1057–1089, 2017.
URL https://doi.org/10.7155/jgaa.00450
Main reference Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, Arnaud Sallaberry: "Path-based supports for hypergraphs", J. Discrete Algorithms, Vol. 14, pp. 248–261, 2012.
URL https://doi.org/10.1016/j.jda.2011.12.009
Main reference Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, Arnaud Sallaberry: "Blocks of Hypergraphs – Applied to Hypergraphs and Outerplanarity", in Proc. of the Combinatorial Algorithms – 21st International Workshop, IWOCA 2010, London, UK, July 26-28, 2010, Revised Selected Papers, Lecture Notes in Computer Science, Vol. 6460, pp. 201–211, Springer, 2010.
URL https://doi.org/10.1007/978-3-642-19222-7_21

A support of a hypergraph is a graph such that each hyperedge induces a connected subgraph. Per definition, a hypergraph is (vertex-)planar if it has a planar support. A c-planar support is a planar support with a planar embedding in which no cycle composed of vertices of a hyperedge S encloses a vertex not in S. In an Euler diagram of a hypergraph each hyperedge S is represented by a simple closed region $R(S)$ bounded by a simple closed curve enclosing exactly the vertices in S. A c-planar support is related to an Euler diagram in which for any two hyperedges S_1, S_2 we have that (a) each connected component of $R(S_1) \cap R(S_2)$ contains a vertex and (b) $R(S_1) \subset R(S_2)$ if $S_1 \subset S_2$.

It can be decided in polynomial time whether a hypergraph H admits a c-planar support if the underlying graph H_2 of hyperedges of size two is already a support. In general, it is NP-complete to decide whether a hypergraph has a c-planar support, even if H_2 is biconnected and induces at most two connected components on each hyperedge. A cactus support is always c-planar. Using a decomposition into blocks, it can be decided in polynomial time whether a hyperedge admits a cactus support.

A support is path-based if each hyperedge contains a spanning path. It can be decided in polynomial time whether a hyperedge has a path-based tree support.

5.4 Short Plane Supports for Spatial Hypergraphs

Martin Nöllenburg (TU Wien, AT)

Joint work of Thom Castermans, Mereke van Garderen, Wouter Meulemans, Martin Nöllenburg, and Xiaoru Yuan
Main reference Thom Castermans, Mereke van Garderen, Wouter Meulemans, Martin Nöllenburg, Xiaoru Yuan: "Short Plane Supports for Spatial Hypergraphs", in Proc. of the Graph Drawing and Network Visualization – 26th International Symposium, GD 2018, Barcelona, Spain, September 26-28, 2018, Proceedings, Lecture Notes in Computer Science, Vol. 11282, pp. 53–66, Springer, 2018.
URL https://doi.org/10.1007/978-3-030-04414-5_4

A spatial hypergraph $H = (V, \mathcal{S})$ is a hypergraph on a set of points $V \subset \mathbb{R}^2$ with $\mathcal{S} \subset 2^V$. A support graph of H is a graph $G = (V, E)$ on the same vertex set V such that for every hyperedge $S \in \mathcal{S}$ the induced graph $G[S]$ is connected. Support graphs are useful geometric structures for various types of set visualizations such as line sets or KelpFusion [1, 2], which enclose or span each hyperedge. We are interested in short supports that use a small amount of ink, that have no edge crossings and that are actually trees. We concentrate on instances with two hyperedges with non-empty intersection.

In the talk I first present a simple sufficient condition for the existence of plane tree supports, based on minimum spanning trees on the non-empty intersection of all hyperedges. However, this idea may lead to plane tree supports being longer by a linear factor than the shortest plane tree support. In fact, it is NP-hard to minimize the total edge length of plane tree supports even if $\mathcal{S}$ contains just two hyperedges. From a practical point of view, I sketch two heuristic algorithms, using either iterated minimum spanning tree computations or local search, possibly relaxing the requirement of planarity or being a tree. An experimental evaluation showed that the local search performs quite well in terms of quality and is still reasonably fast in practice. The extension of this problem to temporal hypergraphs satisfying stability constraints of the support graphs over time is open.

References

1 B. Alper, N. Henry Riche, G. Ramos, and M. Czerwinski. Design study of linesets, a novel set visualization technique. *IEEE Trans. Visualization and Computer Graphics* 17(12):2259–2267, 2011, 10.1109/TVCG.2011.186.

2 W. Meulemans, N. Henry Riche, B. Speckmann, B. Alper, and T. Dwyer. Kelpfusion: A hybrid set visualization technique. *IEEE Trans. Visualization and Computer Graphics* 19(11):1846–1858, 2013, 10.1109/TVCG.2013.76.

6 Panel Discussion

The panel discussion was entitled "Visual Analytics for Sets over Time and Space: What are the burning scientific questions? An interdisciplinary perspective."

6.1 There's a Space for That!

André Skupin (San Diego State University, US)

URL https://datascience.bigknowledge.net/explorer/

The relevance of cartography and geographic thinking extends far beyond the traditional bounds of applying mapping and spatial analysis to study phenomena in geographic space. With some imagination, the elements of any domain can be seen as simultaneously existing in a multitude of spaces, such as attribute space, network space, or knowledge space. Visual analytics then becomes about making any such space accessible to human cognition in order to support more informed decision-making. Once such an overarching spatial viewpoint is in place, any data set can be transformed into engaging artifacts that support research, education, and practice. For example, geographic cells monitored via multi-temporal satellite images can be seen as traversing paths through multi-spectral space. Wildfire events cause rapid shifts of cells, while post-fire recovery carves a slow and steady path over the course of several years. Thanks to dimensionality reduction and visual analytics, we can now SEE this. Meanwhile, natural language processing and machine learning can be leveraged into producing detailed domain base maps, as I presented here for the Data Science & Analytics domain. Individual documents or whole repositories – such as the abstracts of all Dagstuhl Seminars – could now be explored with the ease of everyday mapping interfaces.

6.2 VASet over Time and Space – A GeoVis Perspective

Susanne Bleisch (FH Nordwestschweiz – Muttenz, CH)

Main reference Arzu Çöltekin, Susanne Bleisch, Gennady Andrienko, Jason Dykes: "Persistent challenges in geovisualization – a community perspective". International Journal of Cartography 3(sup1):115–139 (2017).
URL https://doi.org/10.1080/23729333.2017.1302910

The recent paper "Persistent challenges in geovisualization – a community perspective" [1] analyzes and summarizes the input on persistent challenges from four different expert workshops and contrasts them with more top-down research agendas. One of the identified points is also relevant and important for set visualizations – that is, matching data to tasks to visualization types to make it easier to work visually with these data sets.

Additionally, I argue that on the visualization continuum from exploration to communication we too often lean towards optimization of the visualization, which requires knowledge of the task, data insight, the user, etc., and is thus on the communication side. Doing exploration is more difficult as the data and tasks may be ill-defined or, with regard to the tasks, more or less unknown. Similarly with knowing about the users. User knowledge from evaluation is generalized so that we can design for the average user or the majority of users. But our future goal should be personalization. This is ideally not based on preference but rather on performance in terms of gaining knowledge about the data. One potential way to achieve more exploration and more personalization–where in both cases the effectiveness of gaining insights counts–is the combination of the results of exploratory visualizations through new and/or additional forms of visualizations, i.e., by finding the common and constant or the diverging results.

References

1 Çöltekin, Arzu and Bleisch, Susanne and Andrienko, Gennady and Dykes, Jason (2017). Persistent challenges in geovisualization – a community perspective. *International Journal of Cartography*. 3:sup1, 115–139, .

7 Working Groups

7.1 Concentric Set Schematization

Michael Bekos (Universität Tübingen, DE), Fabian Frank (University of Arizona – Tucson, US), Wouter Meulemans (TU Eindhoven, NL), Peter Rodgers (University of Kent – Canterbury, GB), and André Schulz (FernUniversität in Hagen, DE)

Various techniques have been developed to visualize sets, either on a geographically accurate base map or in an abstract non-spatial layout. However, geographic accuracy is often not necessary for overview tasks or tasks focusing on set structures. Yet, completely discarding spatial context may also hide structure or patterns.

Schematic maps have been successful in various applications, such as metro maps, by simplifying and abstracting spatial relations to a minimum functional level, thereby clarifying and emphasizing structure in data while not disregarding (geographic) space.

In this working group, we explore the possibility of computing schematic set visualizations. That is, we are given a set of points in a geographic space, each associated with one or more sets. We want to shift the points to new locations such that we can provide a clear representation for each of the sets; this representation is a geometry connecting (e.g. a tree, cycle or path) or encompassing (e.g. a simple polygon) exactly the points that belong to the set. The main considerations are the extent of the changes we allow to the points, such that we can control for geographic distortion, and the criteria and measures to assess the quality of the resulting set representations.

7.1.1 Sample of Related Work

Set visualization (also known as hypergraph drawing) has received attention in both the visualization and the graph-drawing communities. A recent survey [3] shows variety of visualization techniques; various of these methods target spatial data, e.g. [2, 6, 7, 10]. In the graph-drawing community, most attention has been afforded to hypergraph supports [9] for both fixed and free vertex locations, e.g. [1, 4, 5, 8].

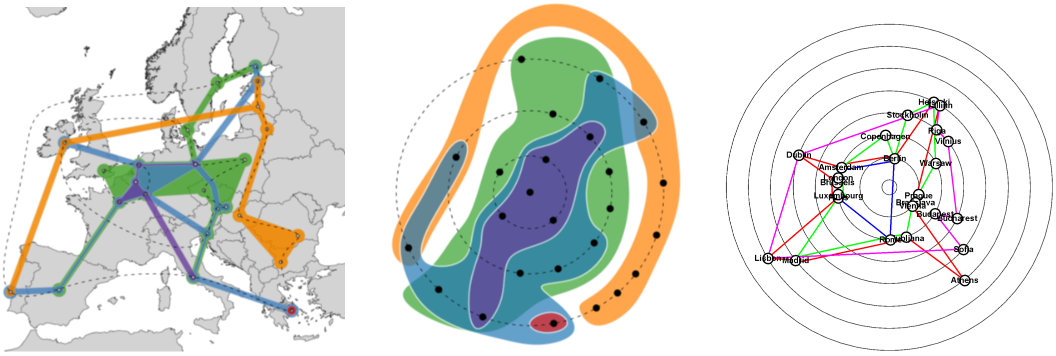

Figure 1 Left: KelpFusion set visualization [10] on a geographically accurate base map. Middle: Manual sketch of a concentric schematic representation of the sets. Right: Preliminary result of our prototype implementation.

7.1.2 Concentric Set Schematization

We study a model where we are given, in addition to the geospatial set system, a set of concentric circles. The goal is to place the given points on the circles such that a clear representation of the set system is given, without distorting the geospatial situation too severely. Through various models of allowed distortion can be considered, we focus on allowing points to move only along the ray through it, originating from the circles' center. For every point an interval of possible circle locations is given with the input. The sets are to be represented as connecting geometries (trees, paths, or cycles). That is, we aim to compute a point placement together with a support of the hypergraph (set system), such that their combination has good quality.

The first question we studied was on computing a good support, measured by the number of intersections the support induces. We found relations to existing work on layered graph drawing and book embeddings. However, the problem as modeled provides different challenges. We were able to show that, given only two circles and a set system of non-intersecting sets, testing whether a planar support exists can be tested easily as follows. We found two configurations, one that implies that a support edge much cross from one circle to the other, and one that implies that a support edge cannot cross. If there are no contradicting configurations, then a planar support exists, by appropriately choosing sides for each point. However, we conjecture that the general version of this problem (intersecting sets, multiple circles) is NP-hard.

The second question we studied assumes that we have a support given (designed or computed by an algorithm), and want to decide for each point on which circle is should be placed. We show that, if we want to minimize the radial change between any points connected by the support, a simple (integer) linear program suffices. We prove that the relaxation has an optimal integer solution and thus the problem can be solved efficiently. In particular, for every LP solution we can remove a set of constraints that are not tight and obtain a LP with the same solution, for which the underlying matrix is totally unimodular. In fact, all vertices of the feasible region induced by the original LP are integral and in bijection to the layer assignments. As a consequence the optimization problem can be solved by greedily improving a layer assignment.

7.1.3 Outlook

These initial findings leave us with a host of interesting questions, both algorithmic in nature and on visual design of concentric set schematization. In particular, we plan to further investigate different models of spatial distortion and criteria for layout quality. We also need to determine how to best route the connecting geometries to obtain an effective design of schematic set representations.

References

1 H. A. Akitaya, M. Löffler, and C. D. Tóth. Multi-colored spanning graphs. In *Graph Drawing and Network Visualization (GD'16)*, LNCS 9801, pages 81–93, 2016.

2 B. Alper, N. Henry Riche, G. Ramos, and M. Czerwinski. Design study of LineSets, a novel set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2259–2267, 2011.

3 B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. The state of the art of set visualization. *Computer Graphics Forum*, 35(1):234–260, 2016.

4 U. Brandes, S. Cornelsen, B. Pampel, and A. Sallaberry. Path-based supports for hypergraphs. *Journal of Discrete Algorithms*, 14:248–261, 2012.

5 T. Castermans, M. van Garderen, W. Meulemans, M. Nöllenburg, and X. Yuan. Short plane supports for spatial hypergraphs. In *Graph Drawing and Network Visualization (GD'18)*, LNCS 11282, pages 1–14, 2018.

6 C. Collins, G. Penn, and S. Carpendale. Bubble Sets: Revealing set relations with isocontours over existing visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009–1016, 2009.

7 K. Dinkla, M. van Kreveld, B. Speckmann, and M. Westenberg. Kelp Diagrams: Point set membership visualization. *Computer Graphics Forum*, 31(3pt1):875–884, 2012.

8 F. Hurtado, M. Korman, M. van Kreveld, M. Löffler, V. Sacristán, A. Shioura, R. I. Silveira, B. Speckmann, and T. Tokuyama. Colored spanning graphs for set visualization. *Computational Geometry: Theory and Applications*, 68:262–276, 2018.

9 D. S. Johnson and H. O. Pollak. Hypergraph planarity and the complexity of drawing Venn diagrams. *Journal of Graph Theory*, 11(3):309–325, 1987.

10 W. Meulemans, N. Henry Riche, B. Speckmann, B. Alper, and T. Dwyer. KelpFusion: A hybrid set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 19(11):1846–1858, 2013.

7.2 From Linear Diagrams to Interval Graphs

Steven Chaplick (Universität Würzburg, DE), Michael Kaufmann (Universität Tübingen, DE), and Stephen G. Kobourov (University of Arizona – Tucson, US)

7.2.1 Introduction

There are many different approaches to visualizing sets. Here we are interested in a particular type of visualization called *linear diagram* [7, 11]. In such visualizations the sets are represented by line segments and the membership of an element in multiple sets is represented by an overlap between the segments in the corresponding sets; see Fig. 2. There is some evidence that linear diagrams are better than Euler diagrams [9].

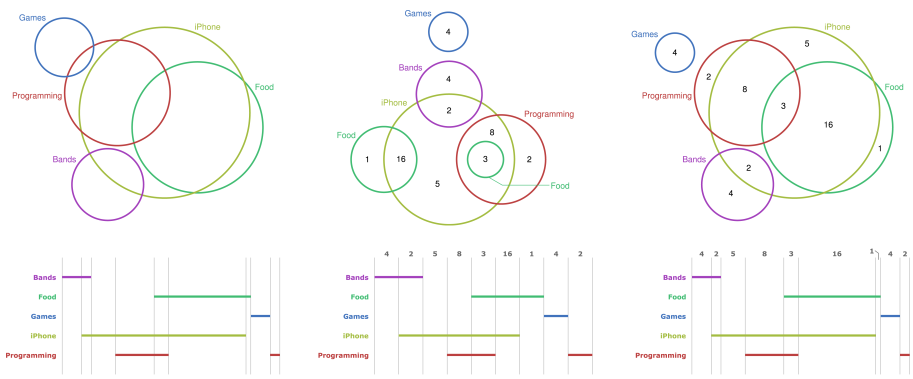

Figure 2 Example of linear diagrams from [9].

We note that linear diagrams are related to the classical interval representation of graphs, originating in combinatorics [6] and genetics [1], and we want to study the connections between the linear diagrams and interval representations. For example, a natural question to ask is given a particular set system, is it possible to represent it so that each set has exactly one segment? The equivalent question is whether a given graph is an interval graph. This problem can be decided in linear time [2, 3]. If such a solution does not exist, then it would be nice to minimize the number of segments per set. However, this problem is NP-hard, even when asking whether 2 segments per set suffice [12].

If we want to visualize linear diagrams over time, there are several parameters to consider:

- do we consider the input given offline (all information given at once) or online (information given incrementally)
- weighted/unweighted (proportional/not) intervals
- vertices have fixed positions over time, or they move
- no splits of intervals are allowed, or we want to minimize the number of splits
- if splits are allowed, we minimized the total number of attributes that get split, or the total number of segments in the representation.

7.2.2 Summary of Discussion

We formalize a *linear diagram* of a set C of n characters $\{c_1, \ldots, c_n\}$ and a set A of k attributes $\{a_1, \ldots, a_k\}$ as a hypergraph $H = (V, E)$ as follows. Let M be the incidence matrix where each row corresponds to a character and each column to an attribute such that the entry $M_{i,j}$ is 1 when character c_i has attribute a_j.

Remarks: across all variants, we should have all characters with the same attribute set consecutive (i.e., consecutive rows). Note that for a single time step we can consider "condensing" twins, e.g., showing the count of the twins instead of each individual.

- Two natural measures of the quality of a linear diagram are the number of segments per attribute and the total number of segments in the whole diagram. It is NP-complete to test if a hypergraph can be realized with 2 segments per attribute (as this is recognition of 2-interval graphs [12]), but it remains open for approximation. On the other hand the status of minimizing the total number of segments seems open.
- As these problems are likely to be computationally difficult, one idea is to employ a simple greedy heuristic to produce some split interval representation, e.g., by iteratively introducing vertices (using the PQ-tree approach) so as to minimize the number of segments for the new vertex. A limitation here is that the first two vertices are certainly never split.

We considered two different models of linear diagrams.

- Model 1: rows are fixed
  - Option 1: suppose we obtain an interval model for each time step (somehow, e.g„ by running the greedy heuristic). If the combined graph is an interval graph, then there is a permutation of the rows consistent with all time steps.
  - Option 2: use heuristic/approximation approach (black-box) to obtain "splits" on all attributes over all time steps. This would provide an interval model for the entire time frame, but it seems unlikely that such diagrams would provide good visualizations (due to having many splits).
- Model 2: rows can move. (when Model 1 fails)
  - (as in Option 1 above): suppose we obtain an interval model for each time step (somehow, eg, by running a greedy heuristic).
  - Limitation: Minimizing the difference between a permutation of time steps $i, i+1$ the problem is NP-hard (as this is related to the tanglegram problem [4]).
  - Potential upside: an efficient algorithm for the related tanglegram problem where one permutation is fixed and the other is flexible (in their case, from a binary tree) [4], seems likely to be able to be generalized to our case when similarly one permutation is fixed and the other should be obtained from an appropriate PQ-tree.

Some further considerations regarding these problems and how to solve them include:

- characters appearing and disappearing over time.
- order of columns, e.g., to maximize overlaps on consecutive attributes or optimize the intervals of the characters (e.g., gaps on their attributes).
- exact algorithms via ILP/SAT formulations, or dynamic programming.

Finally, a further direction in the context of linear diagrams is how they relate to Euler diagrams. In an Euler diagram (e.g., see the top row of Figure 2) the *regions* are the faces of the planarization of the Euler Diagram (made up of closed curves, one for each attribute). If we trace all regions of an Euler Diagram with a closed curve so that this curve visits each region exactly once, then this describes a permutation on the regions. In particular, the number of times the curve visits a each attribute, is the equal to the number of segments in the linear diagram corresponding to the permutation on the regions.

There are some connections to traversing a path in the Euler diagram. But not all sets can be represented by Euler diagrams [8]. There is more background about "well formed Euler diagrams" and what can and cannot be done in these papers [5, 10]. Such connections seem to be an interesting topic for further consideration.

7.2.3 Specific Directions to Study

1. Describe and analyze the greedy heuristic for computing an initial linear diagram, based on inserting the first 2 segments in the PQ-tree without any splits and then splitting as little as possible, given the current PQ-tree. Does this provide any approximation guarantees? (e.g„ to the problem of minimizing the total number of intervals in the linear diagram).
2. In the second model we have a tanglegram variation where the two trees that we want to align are not binary trees but PQ-trees. This means that there are some nodes of high degree and we might need to compute all permutations for their children. This could lead to an efficient algorithm (possibly parameterized by the maximum degree of the PQ-tree).
3. It remains to formalize the precise algorithm to obtain a split interval representation using the PQ-tree approach and to prove that it is *incrementally* optimal.

4. What can we say about the connection to Euler diagrams? For example, if the input set system is nice (well-formed, or whatever property we need), is it true that the best linear diagram corresponds to a Hamiltonian path in the dual of the Euler diagram?

References

1 Seymour Benzer. On the topology of the genetic fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, 45(11):1607–1620, 1959.
2 Kellogg S. Booth and George S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms. *J. Comput. Syst. Sci.*, 13(3):335–379, 1976.
3 Derek G. Corneil, Stephan Olariu, and Lorna Stewart. The LBFS structure and recognition of interval graphs. *SIAM J. Discrete Math.*, 23(4):1905–1953, 2009.
4 Henning Fernau, Michael Kaufmann, and Mathias Poths. Comparing trees via crossing minimization. *J. Comput. Syst. Sci.*, 76(7):593–608, 2010.
5 Jean Flower and John Howse. Generating Euler diagrams. In *International Conference on Theory and Application of Diagrams*, pages 61–75. Springer, 2002.
6 G. Hajós. Über eine Art von Graphen. *Int. Math. Nachr.*, 11:65, 1957.
7 Masood Masoodian and Laura Koivunen. Temporal visualization of sets and their relationships using time-sets. In *22nd International Conference Information Visualisation (IV)*, pages 85–90. IEEE, 2018.
8 Paolo Simonetto and David Auber. Visualise undrawable Euler diagrams. In *12th International Conference Information Visualisation*, pages 594–599. IEEE, 2008. `doi:10.1109/iv.2008.78`
9 Gem Stapleton, Peter Chapman, Peter Rodgers, Anestis Touloumis, Andrew Blake, and Aidan Delaney. The efficacy of Euler diagrams and linear diagrams for visualizing set cardinality using proportions and numbers. *PloS one*, 14(3):e0211234, 2019. `doi:10.1371/journal.pone.0211234`
10 Gem Stapleton, John Howse, and Peter Rodgers. A graph theoretic approach to general Euler diagram drawing. *Theoretical Computer Science*, 411(1):91–112, 2010.
11 Paola Valdivia, Paolo Buono, and Jean-Daniel Fekete. Hypenet: visualizing dynamic hypergraphs. In *EuroVis 2017 – 19th EG/VGC Conference on Visualization*, pages 1–3, 2017.
12 Douglas B. West and David B. Shmoys. Recognizing graphs with fixed interval number is NP-complete. *Discrete Applied Mathematics*, 8(3):295–305, 1984.

7.3 Thread Visualization

William Evans (University of British Columbia – Vancouver, CA), Somayeh Dodge (University of Minnesota – Minneapolis, US), Fabian Frank (University of Arizona – Tucson, US), Panos Giannopoulos (City – University of London, GB), and Giuseppe Liotta (University of Perugia, IT)

We considered the problem of visualizing the communication pattern between concurrent threads in a distributed computation. One particular visualization (see https://bestchai.bitbucket.io/shiviz/) draws vertical lines in 2d to represent individual threads. The y-dimension represents time, increasing from top to bottom. If one thread A sends a message to another thread B at time s, and thread B receives it at time t, then there is a segment connecting the vertical line representing A at y-coordinate s to the vertical line representing

B at y-coordinate t. For complicated communication patterns, these *communication segments* may be hard to distinguish: they may intersect each other and the *thread lines*, nearly overlap, or be nearly vertical. Our working group considered the problem of choosing the position and x-order of the vertical thread lines in order to minimize the number of crossings of the communication segments. We showed that this problem is NP-complete.

Another possible visualization of thread communication is to draw the communication segments vertically between the send and receive time y-coordinates. The thread lines then become y-monotone polygonal chains that connect the endpoints of these communication segments that represent send or receive events experienced by the thread. We considered the problem of choosing the position and x-order of the vertical communication segments in order to minimize the number of crossings in the drawing. We conjecture that this problem in NP-complete as well, but we made some progress on finding an efficient algorithm to test if a plane drawing exists.

7.4 Clustering Colored Points in the Plane

Jan-Henrik Haunert (Universität Bonn, DE), Hugo Akitaya (Tufts University – Medford, US), Sabine Cornelsen (Universität Konstanz, DE), Philipp Kindermann (Universität Würzburg, DE), Tamara Mchedlidze (KIT – Karlsruher Institut für Technologie, DE), Martin Nöllenburg (TU Wien, AT), Yoshio Okamoto (The University of Electro-Communications – Tokyo, JP), and Alexander Wolff (Universität Würzburg, DE)

Analyzing large sets of geographic objects requires visualizations that present the most important spatial patterns in a legible way. Therefore, one often aims to aggregate the given objects to form larger and more abstract entities that can be displayed with low visual complexity. Aggregation is often performed with clustering algorithms that are based on similarity and proximity. Consider, for example, a set of points of interest (POIs) that belong to different categories (e.g., restaurants, shops, etc.). A reasonable clustering approach is to compute a proximity graph of all points (e.g., the Delaunay triangulation), to discard all edges that connect two points of different categories, and to consider each connected component of the remaining graph $G = (V, E)$ as a cluster. Figure 3 shows the result of this procedure for a set of POIs from OpenStreetMap. Similar triangulation-based methods for point-set clustering have been suggested to find groups of animals, e.g., flocks of birds that are relatively near to each other and have the same heading with respect to a few cardinal directions [4].

Although the Delaunay triangulation yields reasonably defined clusters, the obtained solution may not be satisfactory, for example, because the obtained clusters may be considered too small or not compact enough. Therefore, in this report, we introduce more flexibility into the clustering procedure by using substantially denser proximity graphs. In particular, we do not require the graph G used for clustering to be planar and, therefore, need to deal with edge crossings. We still require that each cluster c is connected in G, i.e., G contains a tree T_c spanning the nodes in c. However, we shall not simply report the connected components of G as clusters since, due to the edge crossings in G, it may be difficult to visualize such clusters in a legible way. Instead, we introduce the following basic optimization problem to find large non-overlapping clusters.

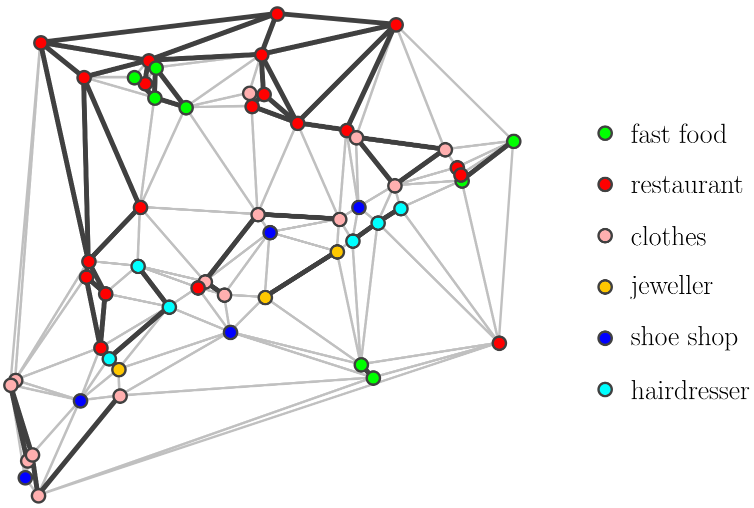

Figure 3 A set of points of interest from OpenStreetMap with its Delaunay triangulation. Edges connecting two points of the same category are displayed fat. These edges induce a subgraph $G = (V, E)$ of the Delaunay triangulation whose connected components can be considered as clusters.

▶ **Cluster Minimization.** Let $G = (V, E)$ be a geometric graph that may contain edge crossings. Find a subgraph H of G with node set V, i.e., a graph $H = (V, E')$ with $E' \subseteq E$, such that no two edges of E' cross and the number of connected components of H is as small as possible.

Note that we do not require the nodes of G to belong to any category. We will, however, focus on special cases of the problem that require categorized points as input. Generally, we refer to the categories as *colors*. Unless stated otherwise, we require that each point is assigned to exactly one of k colors. Consequently, we refer to G as *k-partitioned*. An edge connecting two points of the same color is called *colored* and its color equals the common color of its two end points; other edges are called *uncolored*. A crossing of two colored edges is called *monochromatic* if both edges have the same color and, otherwise, *bichromatic*. *Planarizing* a monochromatic crossing of two edges $e = \{u, v\}$ and $f = \{p, q\}$ means introducing a new node x of the same color as the four involved nodes at the intersection of e and f and replacing these edges with edges $\{u, x\}$, $\{v, x\}$, $\{p, x\}$, and $\{q, x\}$. A graph is *1-planar* if each edge is crossed by at most one other edge. With these definitions, we are ready to define special cases of Cluster Minimization for our further investigations:

▶ **Cluster Minimization in 1-Planar Graphs.** G is obtained from a k-partitioned, 1-planar graph that contains only colored edges, by planarizing monochromatic crossings. The problem is to solve Cluster Minimization for G.

▶ **Cluster Minimization in Complete Graphs.** G is k-partitioned and complete in the sense that every two points of equal color are connected with an edge and there is no uncolored edge. The problem is to solve Cluster Minimization for G.

Cluster Minimization in 1-Planar Graphs arises in the situation that G is constructed by computing a 1-planar proximity graph of the given points, reducing it to its colored edges, and planarizing its monochromatic crossings. For example, one may define the edge set of the proximity graph as the union of the edge sets of all order-1 Delaunay triangulations of the point set. The resulting proximity graph is 1-planar [3] and, due to its close relationship with the Delaunay triangulation, reflects the proximity relationships of the points reasonably well. On the other hand, since this graph has substantially more edges than the Delaunay triangulation, an optimal solution of Cluster Minimization may contain fewer and larger clusters than the approach based on the Delaunay triangulation illustrated in Figure 3. The idea behind Cluster Minimization in Complete Graphs is to define the graph G as large as possible and, thereby, to place even more emphasis on creating few and large clusters.

Finally, we introduce a problem that is of relevance if one wishes to place more emphasis on creating compact clusters.

▶ **Edge Maximization.** Let $G = (V, E)$ be a k-partitioned geometric graph that contains only colored edges and that may contain monochromatic as well as bichromatic edge crossings. Find a subgraph H of G with node set V, i.e., a graph $H = (V, E')$ with $E' \subseteq E$, such that $|E'|$ is as large as possible and H contains no bichromatic edge crossing.

Again, G may be obtained by computing some proximity graph and reducing it to its colored edges. The idea behind maximizing the number of edges of H is that a set of nodes that is densely connected in G constitutes a compact cluster. Bichromatic edge crossings in the output graph H are forbidden to avoid overlapping clusters of different colors. Monochromatic edge crossings, on the other hand, are allowed since they could be planarized before computing the output clusters as the connected components of H.

7.4.1 Results

We have obtained the following main results for the problems defined above:

- Cluster Minimization in 1-Planar Graphs can be solved efficiently with a simple greedy algorithm.
- Cluster Minimization in Complete Graphs can be solved efficiently if there are only $k = 2$ different colors, by adapting an algorithm by Bereg et al. [1]. For $k = 3$, however, it is NP-hard, which we showed by reduction from Tree-Residue Vertex-Breaking [2].
- For general Cluster Minimization, we devised an integer linear program and a heuristic based on the greedy algorithm for the 1-planar case.
- Edge Maximization can be solved efficiently for $k = 2$ using an algorithm for minimum vertex cover in bipartite graphs. For arbitrary k and 1-planar graphs, the problem is trivial since for every edge crossing one can arbitrarily decide which of the two involved edges to select.

7.4.2 Future Work

We plan to continue our research with respect to both (a) a further theoretical study of open problems and (b) an evaluation of our clustering approach based on implementations of our algorithms and experiments with real-world data. We already started to develop ideas for the case that each point has more than one color and we were able to generalize some of our algorithms for this case. Moreover, we have first ideas on how to deal with a dynamic situation in which the points move or change their colors over time and the aim is to compute a dynamic visualization while considering the stability of the visualization as an additional optimization criterion.

References

1 S. Bereg, M. Jiang, B. Yang, and B. Zhu. On the red/blue spanning tree problem. *Theoret. Comput. Sci.*, 412(23):2459–2467, 2011. doi:10.1016/j.tcs.2010.10.038.

2 E. D. Demaine and M. Rudoy. Tree-Residue Vertex-Breaking: A new tool for proving hardness. In D. Eppstein, editor, *Proc. 16th Scandinavian Symp. Algorithm Theory (SWAT'18)*,

volume 101 of *LIPIcs*, pages 32:1–32:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.SWAT.2018.32.

3 J. Gudmundsson, M. Hammar, and M. van Kreveld. Higher order Delaunay triangulations. *Comput. Geom. Theory Appl.*, 23:85–98, 2002. doi:10.1016/S0925-7721(01)00027-X.

4 P. Laube, M. van Kreveld, and S. Imfeld. Finding REMO – Detecting relative motion patterns in geospatial lifelines. In *Developments in Spatial Data Handling – Proc. 11th Int. Symp. Spatial Data Handling (SDH'05)*, pages 201–215. Springer-Verlag, Berlin, Germany, 2005. doi:10.1007/3-540-26772-7_16.

7.5 Flexible Visualization of Sets over Time and Space

Christian Tominski (Universität Rostock, DE), Gennady Andrienko (Fraunhofer IAIS – Sankt Augustin, DE), Natalia V. Andrienko (Fraunhofer IAIS – Sankt Augustin, DE), Susanne Bleisch (FH Nordwestschweiz – Muttenz, CH), Sara Irina Fabrikant (Universität Zürich, CH), Eva Mayr (Donau-Universität Krems, AT), Silvia Miksch (TU Wien, AT), Margit Pohl (TU Wien, AT), and André Skupin (San Diego State University, US)

Sets with references to time and space are difficult to visualize. The reason is that multiple aspects of the data must be communicated to the user. In the first place, the set characteristics need to be encoded visually. Moreover, data attributes associated with the set members are important information to be visualized. On top of that, the temporal and spatial frames of reference need to be displayed. Uncertainty of the data might also play a role. Encoding all of these aspects into a single visual representation is typically impractical because the result would likely be complicated to interpret. An alternative is to use multiple views where each view focuses on selected aspects of the data. Yet, connecting findings made in one view to findings made in another view can be a considerable effort. On an elementary level, the user needs to understand how a data object in one view, i.e., in one reference system, relates to another perspective in another view or reference system.

Given the two extremes of fully integrating all aspects in a single representation and separating selected aspects into multiple representations, the working group discussed an alternative option in between, which we call *flexible visualization.* The idea is to bring the two extremes of integration and separation closer together by means of animated transitions. The starting point is to have visual representations that show the data with selected aspects being prioritized and other aspects being attenuated or omitted. For example, a set of movement trajectories is shown on a 2D map to prioritize the spatial aspect of the data. Another view might show the same data as stacked 3D bands above a map to better reveal the data attributes along individual trajectories. Yet another view might show the data as horizontal 2D bands to emphasize the temporal aspect of the data. Now, the core idea of flexible visualization is to have smooth transitions that transform one view into another, rather than showing them as multiple views. That said, flexible visualization aims to balance visual complexity and interaction while providing opportunities to see data and patterns from multiple perspectives.

There are already existing approaches that implement smooth transitions to flexibly animate between views. We recognized the need for a systematic approach to categorizing flexible visualizations in order to gain a better understanding of the potential and limitations

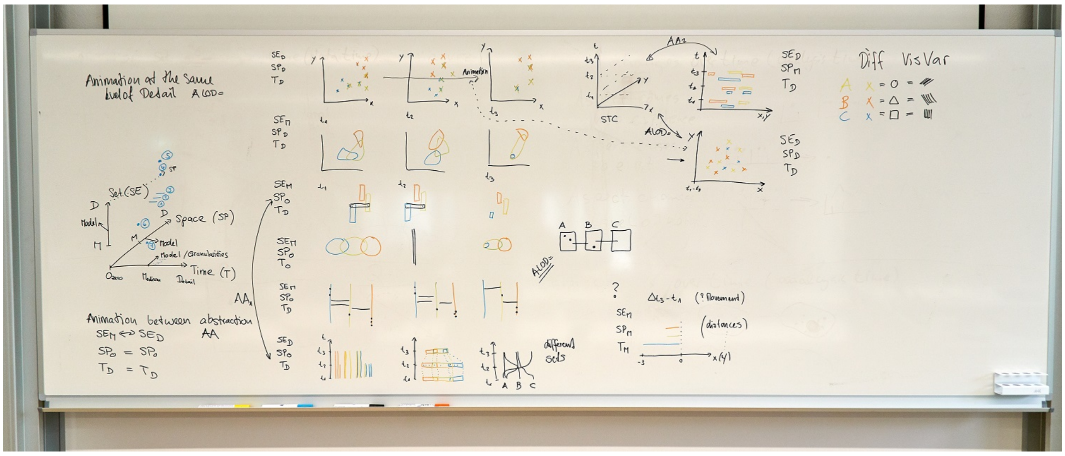

Figure 4 Sketches to systematize flexible visualization.

of augmenting the visual analysis by transitions between discrete visual states. The working group split up into two subgroups to discuss in detail several questions related to flexible visualization, including:

1. Conceptual and technical aspects
   - What are the requirements for flexible visualization?
   - What principle transitions between views are possible and make sense?
   - What topology might flexible visualizations exhibit?
   - Where can smooth transitions operate, in data space or in view space?
   - How can flexible visualization be implemented?
2. Human aspects
   - What are perceptual and cognitive constraints?
   - How should an animated transition be designed?
   - What is the role of interactive user control?
   - Where is the sweet spot between abrupt change and very smooth transitions?
   - Does flexible visualization scale to very large data?

The working group developed first sketches to systematize flexible visualization. Figure 4 shows an example. A first draft of a research publication has been prepared. The goal of this publication is to characterize flexible visualization comprehensively as a viable approach to enhance the visual analysis of complex multi-aspect sets.

Participants

- Hugo Akitaya
Tufts University – Medford, US
- Gennady Andrienko
Fraunhofer IAIS – Sankt Augustin, DE
- Natalia V. Andrienko
Fraunhofer IAIS – Sankt Augustin, DE
- Michael Bekos
Universität Tübingen, DE
- Susanne Bleisch
FH Nordwestschweiz – Muttenz, CH
- Steven Chaplick
Universität Würzburg, DE
- Sabine Cornelsen
Universität Konstanz, DE
- Somayeh Dodge
University of Minnesota – Minneapolis, US
- William Evans
University of British Columbia – Vancouver, CA
- Sara Irina Fabrikant
Universität Zürich, CH
- Fabian Frank
University of Arizona – Tucson, US
- Panos Giannopoulos
City – University of London, GB
- Jan-Henrik Haunert
Universität Bonn, DE
- Michael Kaufmann
Universität Tübingen, DE
- Philipp Kindermann
Universität Würzburg, DE
- Stephen G. Kobourov
University of Arizona – Tucson, US
- Giuseppe Liotta
University of Perugia, IT
- Eva Mayr
Donau-Universität Krems, AT
- Tamara Mchedlidze
KIT – Karlsruher Institut für Technologie, DE
- Wouter Meulemans
TU Eindhoven, NL
- Silvia Miksch
TU Wien, AT
- Martin Nöllenburg
TU Wien, AT
- Yoshio Okamoto
The University of Electro-Communications – Tokyo, JP
- Margit Pohl
TU Wien, AT
- Peter Rodgers
University of Kent – Canterbury, GB
- André Schulz
FernUniversität in Hagen, DE
- André Skupin
San Diego State University, US
- Christian Tominski
Universität Rostock, DE
- Alexander Wolff
Universität Würzburg, DE