Towards a Survey on Static and Dynamic Hypergraph Visualizations

Maximilian T. Fischer † Alexander Frings ‡ Daniel A. Keim § Daniel Seebacher ¶

University of Konstanz, Germany

ABSTRACT

Leveraging hypergraph structures to model advanced processes has gained much attention over the last few years in many areas, ranging from protein-interaction in computational biology to image retrieval using machine learning. Hypergraph models can provide a more accurate representation of the underlying processes while reducing the overall number of links compared to regular representations. However, interactive visualization methods for hypergraphs and hypergraph-based models have rarely been explored or systematically analyzed. This paper reviews the existing research landscape for hypergraph and hypergraph model visualizations and assesses the currently employed techniques. We provide an overview and a categorization of proposed approaches, focusing on performance, scalability, interaction support, successful evaluation, and the ability to represent different underlying data structures, including a recent demand for a temporal representation of interaction networks and their improvements beyond graph-based methods. Lastly, we discuss the strengths and weaknesses of the approaches and give an insight into the future challenges arising in this emerging research field.

Keywords: Hypergraphs, hypergraph model, temporal, visualization, visual analytics, survey

1 INTRODUCTION

The modeling of complex network structures has risen to prominence over the last few years in many areas, ranging from protein-interaction in computational biology to image retrieval using machine learning. Hypergraph models can provide a more accurate representation of the underlying processes while reducing the overall number of links compared to regular representations. While modeling problems as hypergraphs can provide benefits, the visualization of the hypergraphs themselves or even the hypergraph models are not trivial tasks. Mathematically, an undirected hypergraph $H = (V, E)$ is defined as an ordered pair, where $V = \{v_1, \ldots, v_n\}$ represents the $n$ vertices (hypernodes) and subsets of these vertices form the multi-set $E = \{e_1, \ldots, e_m\}$, constituting the $m$ distinct hyperedges. We do not explicitly differentiate between classes of hyperedges, i.e., we do not consider heterogeneous hypergraphs separately. Traditionally, the visualization of such hypergraphs uses Venn diagrams or Euler diagrams. As part of the field of graph drawing, their traditional representation has been described in detail [8, 21, 28]. However, as these traditional representations often use color to distinguish between hyperedges, their scalability is severely limited. While hypergraph visualizations can share some characteristics with multivariate graphs [43], the strategies are not always suitable for hypergraphs, which need a more efficient way of depiction [17]. The problem of hypergraph visualization has even become more complex by the recent interest in the evolution of such relations. These temporal (or dynamic) hypergraphs additionally encode time for direct comparisons of network states in different stages and require new visualization approaches for an effective representation. However, more advanced hypergraph visualizations are relatively new and not systematically explored, lacking detailed comparisons between approaches so far.

To address these issues, we present a survey of hypergraph visualizations, making the following contributions:

- a systematic literature review of existing approaches for static and dynamic hypergraph (model) visualization.
- a methodology for comparison criteria between hypergraphs, critically assessing the different approaches.

This survey aims to provide insights into the existing research landscape of generic hypergraph visualizations and identify promising research gaps for future work.

2 RELATED WORK

Work on more advanced hypergraph visualizations has only recently gained traction. Consequently, most surveys cover hypergraph visualizations only as a sub-part – if at all – and rarely provide an extensive overview of the existing methodologies with respective advantages and drawbacks. Therefore, to compile a list of existing approaches, we also consider more generic surveys on set- and (regular) graph visualizations [10, 26, 40, 41], which focus on relational aspects, as some aspects are comparable.

Alsallakh et al. provide a comprehensive overview [2] of set visualizations, which can be regarded as a specific representation of hypergraphs. Many of these approaches are different or enhanced variants of Euler and Venn diagrams like BubbleSets [11], but it also contains several different node-link and matrix-based approaches. However, not all apply to generic hypergraphs, as they are sometimes very domain-specific. In this survey, we only cover generic hypergraph visualizations.

Vehlow et al. analyzed group visualizations using different regular graph representations in their survey [40]. They propose a taxonomy for visualization techniques separated into visual node attributes, juxtaposed, superimposed, and embedded visual styles. This survey forms the groundwork for contrasting group structures and helps in...
categorizing static hypergraph approaches, covering several basic representations but does not include more recent approaches.

When shifting the focus on dynamic graphs, one of the latest surveys [6] on dynamic graph visualizations was compiled by Beck et al. in 2017, focusing on group structures in regular graphs and their relationship over time. They differentiate between animated node-link diagrams and timeline structures to convey the dynamic relationships visually. The authors also state that approaches utilizing timelines are becoming more common in the literature.

For a more set theory-focused approach, the survey [10] by Chen et al. focuses on exploring association relationships in graphs. The authors propose a pipeline for visual analysis of associated data and summarized many different graph representations for large relationships data sets. Furthermore, they discussed the advantages and disadvantages of some visualization types, including interaction combined with simplification methods. As part of their work, they list two visualization methods for hypergraphs, one being the radial layout by Kerren et al. [23], the other being a fixed-node layout presented by Xia et al. [44]. However, this survey seems incomplete and somewhat outdated regarding hypergraph visualizations, as it did not cover some available approaches.

Also, as new hypergraph visualization methods have been proposed since these works have been released and the lack of an extensive survey for hypergraphs, it becomes apparent that there is a gap for an overview of such techniques specifically. Although many proposed pipelines and taxonomies can also be applied to hypergraph models, they do not provide in-depth comparison grounds for them and their dynamic counterparts. For example, Vehlow et al. [40] covered most basic representations, but recent hybrids [16, 37, 39] are partly incompatible with their differentiation criteria.

Recently Valdivia et al. begin to tackle the research gap by proposing PAOVis [39], thereby providing the “first [...] highly readable representation of dynamic hypergraphs”. It features a timeline-like view and linearly orders the intervals to directly compare the hyperedges and the whole structure side-by-side. This approach, by design, is especially suited for a medium-sized (less than 100) number of relatively small hyperedges, favoring comparability over time instead of at one time. For dynamic hypergraphs, a hybrid visualization [37] was proposed by Streeb one year earlier. They use arrow-glyphs for time-frames encoding the network changes.

The idea was improved and then generalized to arbitrary dynamic hypergraph models by Fischer et al. [16]. Hyper-Matrix consists of a geometric deep learning model architecture and presents a novel multi-level, matrix-based hypergraph visualization.

3 Methodology
To analyze state-of-the-art approaches on hypergraph visualizations, we conducted a survey of relevant literature before collecting further approaches via cross-references. We target the subset of visualization literature that applies to hypergraphs and hypergraph models, and those techniques that focus on representing hypergraph structures and how they are reflected and realized in the visualization.

3.1 Source and Selection Methodology
Orienting ourselves on semi-automated structured literature surveys, such as the one of Sacha et al. [35], we initialize the paper collection process using a keyword-based search for “hypergraph”, and common variations like “hyper-graph, hyper graph, ...”. To further focus our survey, we consider only approaches from the following high-quality journals and conferences since 2000:

- IEEE Transactions of Visualization and Computer Graphics (TVCG, incl. IEEE VIS proceedings)
- Computer Graphics Forum (CGF, incl. EuroVis and EuroVA)

For the actual paper selection methodology, we follow a three-step approach (see also Figure 1). Automated filtering resulted in 5 (TVCG) and 21 (CGF) approaches; manual filtering reduced this first to 3 and 10, while in a second iteration after careful considerations, to 3 and 6, i.e., nine approaches. Using cross-references and survey data, we identified an additional four relevant approaches, leading to 14 approaches in total, which we compare using the following comparison framework.

3.2 Comparison Framework
We present seven comparison criteria covering some aspects common to graphs and some favoring aspects specifically adapted to hypergraphs. We curated this selection based on an adapted combination of criteria taken from existing surveys [6, 40] as well as aspects discussed in the literature [16, 37, 39], and complemented them by additional distinguishing criteria.

Representation Method The first criterion provides a rough separation into visualization types, thereby deciding on the fundamental technique. The existing approaches can be classified as either node-link diagrams, timeline-based techniques, or matrix-based approaches.

Scalability The supported data-set sizes and, accordingly, the scalability of the approaches (nodes and hyperedges) present a significant factor. Diving the capabilities of existing approaches in five distinguishable groups, we distinguish between very low scalability (less than ten nodes or hyperedges), low scalability (between 10 and 50), medium scalability (between 50 to 200), high scalability (between 200 and 1000), and very high scalability (more than 1000). Especially node-link-diagrams with too many connections are prone to clutter. Many basic representations are only applicable to few vertices and hyperedges, whereas different techniques perform better.

Static vs. Dynamic This criterion is concerned with the support of dynamic (i.e., temporal) data in contrast to static data. Analyzing trends requires different time steps to be comparable in the visualization. The methods range from a simple timeline to a direct comparison of single components by highlighting different stages. Nevertheless, approaches supporting dynamic hypergraphs are still limited and only recently began to emerge.

Interactivity This criterion describes the support of interactivity by visualizations compared to a static, non-interactive visual representation. It describes the ability to modify the model or the resulting representation for analysis purposes. Besides highlighting, sorting or filtering can enable the analysis of larger graphs. For example, even Venn diagrams would benefit from interactivity to reduce ambiguities. Other visualization methods may depend on a multi-level structure to represent the entire hypergraph coherently.

Tasks & Use Cases This criterion describes if the approaches present tasks in the application domain and their specific uses. For example, temporal networks of cell interactions may use comparisons of the sub-graph as small multiples to outline its development.

Evaluation The last comparison criterion concerns evaluations, case studies, and user studies conducted on the different visualization methods. Only a few approaches provide no evaluation at all. However, some evaluations are performed as case studies, for example, by comparing the scalability to other representation types like parallel coordinate plots to demonstrate hypergraph model advantages. In contrast, others conduct both a full user study with participants, often domain experts, concerned with performing analytical tasks, but rarely a quantitative study.
4 Literature Survey

For the following survey, we group the individual techniques based on the representation method and, secondly, their publication date.

4.1 Node-Link-based Approaches

Bubble Sets Collins et al. [11] describe hulls with contouring (marching squares) encompassing connected elements. Interaction is provided through movement and the addition or deletion of nodes.

Software Artifact Hypergraph Visualization Kapec et al. presented a visualization of software functions in a visual programming environment, using hypergraphs as a basis [22]. Source code and calling sequences are modeled as hypergraph-connections using a force-directed 3D layout with spheres and directed links, while color is used to distinguish callers and callees.

EGAN Hypergraph Visualization Paquette et al. [32] propose an extensible visualization method for Exploratory Gene Association Networks (EGAN) to sort and classify gene lists. It uses a node-link-based approach but adds a separate association meta-node that connects to each related entity, imitating the concept of hyperedges, thereby improving scalability compared to color-based methods and being similar to node-link representations.

Kelp Diagram Dinkla et al. [14] extend on the concept of colored-hulls and bubble-maps by using methods like overlapping lines, node color, or varying link sizes. It is primarily suited for static and fixed, geospatial applications, not providing interactions, and limited in scalability. A successor, KelpFusion [27], improves performance while reducing clutter, combining hull and linear set representation.

Radial Representation Kerren et al. [23] present a different layout where hyperedges are circular dotted lines around centered nodes, thereby eliminating overlaps and slightly increasing the scalability. The visualization offers rich interaction support from highlighting groupings, filtering, and link modification and was evaluated in a small user study. The authors present possible extensions, for example, labeling, graph comparison, or an extended study.

Visual Analysis of Set Relations in a Graph Xu et al. [45] present a glyph-based overlay in node-link diagrams to study the effect between shared set relations and node distances in a graph. However, scalability is very limited and interaction concepts are only discussed as future work.

Extra-Node Representation Ouvard et al. [31] present an improved method for node-link conversion of hypergraphs by adding artificial extra-nodes which combine and merge multiple edge connections, retaining more information and inducing less clutter, which is shown qualitatively and quantitatively.

SimpleHypergraphs.jl SimpleHypergraph.jl is a julia library by Antelmi et al. [3], offering established interactive hypergraph visualizations using a combination of D3 and Python. The visualization approaches are similar to EGAN [32], but representing hyperedges as sub-graphs, but also support a convex hull approach similar to a Venn diagram, and support some limited interactivity.

MetroSets The MetroSets approach by Jacobsen et al. [18] represents geospatial hypergraph data, and its primary application is the visualization of train networks. The drawbacks of the model are its enforced octolinearity and unavoidable crossings leading to clutter. Interaction support ranges from highlighting to filtering different sets, but not for node repositioning. As one of a few approaches, it supports temporal data mapped to the x-axis, representing hyperedges at different time-steps.

4.2 Timeline-based Approach

TimeSets Nguyen et al. [29] leverage a stacked timeline view with hyperedges encoded via color, similar to KelpFusion. It offers different interaction techniques, a strong case as well as user study, but suffers from poor scalability without aggregating.

PAOHvis Valdivia et al. introduce PAOHvis [39], an ordered timeline layout to represent hyperedges. It is based on earlier work [38], follows a grid timeline pattern and nodes on the y-axis. Hyperedges are represented through vertical lines connecting multiple nodes, where drip dots can mitigate overflowing. The layout optimizes space usage and allows for parallel vertical hyperedges to split into multiple groups, offering rich filtering and interaction controls like timeflow comparison.

Set Streams Agarwal et al. [1] leverage branching and merging streams in a timeline view to represent dynamic set membership. Different set operations can be applied through queries, and linked lists offer more detailed information, while the scalability extends to several hundred elements.

4.3 Matrix-based Approach

Visual Analytic Framework for Dynamic Hypergraphs Streeb et al. [47] propose a prototype for temporal hypergraph analysis by leveraging a glyph-based matrix view. It supports temporal data through time glyphs but may require interaction to retrieve information, and its scalability is limited.

Hyper-Matrix The latest visualization method was proposed by Fischer et al. [16] and leverages a six-level matrix-based representation for visualizing dynamic hypergraph structures. A geometric deep learning model interfaces with a hypergraph model, and the interface supports multiple filtering, grouping, and interaction mechanism, including matrix-reordering, hierarchical grouping, and the direct integration of domain knowledge to influence the underlying machine learning model.

4.4 Further Methods

During our initial search for related work (searching for hypergraph on Google Scholar), we discovered some related approaches in different disciplines. In the following, we shortly reference approaches where the focus is a task-dependent problem related to a specific hypergraph representation in a different domain and not focused on visualization research, and therefore not full-filling the inclusion criteria of this survey for a generic and generalizable approach.

Early works by Jin et al. [20] present combinations of hypergraph and matrix visualizations through hyperrectangles. Arafat et al. [5] adapt force-directed placement for increased scalability of hull drawing, while Cooper et al. [12] using hypergraphs for topic dependency sketching through varying thickness and colored lines for the association. Cromar et al. [13] visualize protein-protein interactions using a combination of Venn and node-link diagrams, while Liu et al. [25] uses clustered glyphs layouts for uncertainty modeling.

5 Comparison and Discussion

In this survey, we studied currently proposed visualization techniques for hypergraphs and hypergraph models. Table 1 provides an overview of the different approaches and highlights commonalities as well as differences. The presented approaches utilize different interactive features to extend their techniques. In general, node-link-based approaches can only support edge counts ranging from few (Venn diagram) up to several dozens, depending on the technique. In comparison to this, timeline- and matrix-based approaches can support medium- and large-sized data-sets with several hundred nodes and hyperedges. Additionally, there is a clear evolution, as dynamic support and advanced interactivity are featured more prominently in more recent techniques. Interestingly, user studies have also become more common over time. Finally, the aim mainly revolves around improving performance compared to non-hypergraph visualizations.

Most hypergraph visualizations are based on node-link diagrams (or variants thereof) and use regular graph drawing with special nodes or hulls. The work by Kapec [22] sticks out as an early technique providing an extended, three-dimensional, and scalable model compared to its contemporaries. However, all styles use different colors or shapes on their sub-graphs as visual variables to encode node affiliations. Consequently, some methods have difficulties
presenting an accurate picture of the structure as they rely on overlapping projections or might lose information about hyperedges in the drawing process. The interactive adaptations boil down to manipulating positions or viewpoints of the model. The main drawback is the scalability, as many methods become illegible quickly.

The more advanced node-link-based visualizations offer more accurate, powerful, and scalable methods, encoding the hypergraph (models) using an adaption of preexisting techniques. For example, all of them use regular edges. However, they should be interpreted as forming hyperedges through the nodes, sometimes with additional visuals for improved clarity. The exception is the Kelp-diagram approach [14], as it specializes in fixed point visualizations useful for marking geospatial information. However, this heavily limits its applicability and scalability when compared to the other two approaches. For example, the radial representation [2] can visualize more hyperedges side-by-side while the MetroSets [18] deal with overlap when excessively extended. They provide basic-level interactions like highlighting, necessary for interpreting larger data sets. Still, only the MetroSets [18] features labels applied to both nodes and hyperedges. Further, they are also the only technique that incorporates support for dynamic data, while radial representations may only partly support it. Nevertheless, the amount of information this style can display is increased compared to the earlier approaches.

The most common tasks are the search for nodes in the visualization [14][16][18][22][23][29][32][37][39][45] and the determination of communities/memberships and connected components [1][14][16][18][22][23][29][32][37][39][45], sometimes to determine their size [29]. Other common techniques are membership queries [1][14][16][18][22][23][29][32][37][39][45] and the tracing of connections [1][11][16][18][22][23][29][32][37][39]. The latter is sometimes extended to determine implicit shares and overlap [45] and random walks [4]. Less frequent tasks are filtering based on external factors [1][16][23], evaluating the readability [29], classification [1], and change detection in models [16].

The matrix-based visualizations feature more scalable techniques to display large correlated data-sets. They additionally provide extensive interactions to higher degrees, needed to work with the visualizations in realistic scenarios. The PAOHvis approach [39] uses a column-wise timeline representation of hyperedges while the other two methods compare the nodes and their affiliations directly in the cells. To achieve this, they use the nodes as coordinates while different layers of time or tags represent the underlying data. Still, Streeb et al.’s technique [37] is the only one to render a classic node-link-based model as part of their visualization. The latest technique, Hyper-Matrix [16], is equally easy to handle but focuses more on connection development for predicting future affiliations and extends the concept to a generic blueprint for hypergraph model visualization through a new architecture and novel visualization technique.

6 Future Directions

Even with recent advances, opportunities for future work remain. The node-link-based approaches provide a very limited scalability, and even timeline [1] and matrix [16] representations do not support more than around thousand entities. Solving these issues requires novel ideas like extra-nodes [31], aggregating and subsetting, as well as dense, domain-specific representations. Support for dynamic hypergraphs is offered only by six approaches and is often optimized for specific use cases. Specifically, the support for dynamic hypergraphs with many time steps is very limited. One could, for example, explore how novel concepts in dynamic networks visualizations beyond animations [9] are applicable to hypergraphs. Further, there is no established benchmark dataset for hypergraph visualizations, no established performance metrics, and only a limited discussion [16][39] on specific tasks.

7 Conclusion

In the last decade, problem modeling through hypergraphs has gained much attention. While some visualization methods have been proposed, there is a noticeable and increasing gap between applied hypergraph research and their visualization. By surveying the existing approaches for hypergraph (model) visualizations, we aim to structure the research space. We first establish the need for such a survey by analyzing existing literature before defining a reproducible paper selection process. Then, we systematically structure comparison criteria before presenting the techniques individually. We discuss the particularities of each technique individually, before classifying it, and finally discuss in detail the observations in context. We find that many representations do not capture the full potential of hypergraphs and are limited in scalability, interactivity, or the support of dynamic hypergraphs. The three most promising and generic techniques, PAOHvis [39], Set Streams [1], and Hyper-Matrix [16] are also the most recent ones, but all approaches need further refinement. This supplements our observations and highlights the specific opportunities for future work we have identified.

By filling this gap with an overview of hypergraph visualization methods, we aim to provide researchers with a standard reference, promote areas for future work, and set the baseline for a more in-depth survey on hypergraph visualizations.

Acknowledgments

The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany (BMBF) in the framework of PEGASUS under the program “Forschung für die zivile Sicherheit 2018 – 2023” and its announcement “Zivile Sicherheit - Schutz vor organisierter Kriminalität II”.
