

Received October 31, 2020, accepted December 4, 2020, date of publication December 28, 2020, date of current version January 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3047616

The State of the Art in Empirical User Evaluation of Graph Visualizations

MICHAEL BURCH¹, WEIDONG HUANG², MATHEW WAKEFIELD³,
HELEN C. PURCHASE⁴, DANIEL WEISKOPF⁵, AND JIE HUA⁶

¹Institute for Data Analytics, Visualization, and Simulation (DAViS), University of Applied Sciences, 7000 Chur, Switzerland

²TD School, University of Technology Sydney, Ultimo, NSW 2007, Australia

³School of Software and Electrical Engineering, Swinburne University of Technology, Hawthorn, VIC 3122, Australia

⁴School of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK

⁵Visualization Research Center (VISUS), University of Stuttgart, 70569 Stuttgart, Germany

⁶Faculty of Engineering and Information Technology, Shaoyang University, Shaoyang 422000, China

Corresponding author: Jie Hua (2013021250@m.scnu.edu.cn)

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG), German Research Foundation, under Project 251654672-TRR 161.

ABSTRACT While graph drawing focuses more on the aesthetic representation of node-link diagrams, graph visualization takes into account other visual metaphors making them useful for graph exploration tasks in information visualization and visual analytics. Although there are aesthetic graph drawing criteria that describe how a graph should be presented to make it faster and more reliably explorable, many controlled and uncontrolled empirical user studies flourished over the past years. The goal of them is to uncover how well the human user performs graph-specific tasks, in many cases compared to previously designed graph visualizations. Due to the fact that many parameters in a graph dataset as well as the visual representation of them might be varied and many user studies have been conducted in this space, a state-of-the-art survey is needed to understand evaluation results and findings to inform the future design, research, and application of graph visualizations. In this article, we classify the present literature on the topmost level into graph interpretation, graph memorability, and graph creation where the users with their tasks stand in focus of the evaluation, not the computational aspects. As another outcome of this work, we identify the white spots in this field and sketch ideas for future research directions.

INDEX TERMS Aesthetic criteria, algorithms, evaluation, graph drawing, graph visualization, user studies.

I. INTRODUCTION

Graph visualization and graph drawing have become frequently studied fields of research [1], [2]. Novel techniques are designed and implemented, as well as adapted for bigger dataset scenarios. One reason for the increased focus in these fields is the variety of applications that must deal with relational information such as coupling data in software development, protein interactions in bioinformatics, contacts between people in social networking, or schematic maps in the field of cartography or public transportation.

In some scenarios, it is not just the relational information given in a dataset that needs to be visualized, but also the weight, multitude, or direction of relations. Relations may also carry additional attributes in the case of multivariate datasets. Graph vertices can also hold additional properties,

for example, they might be hierarchically structured in some way, which is of special interest for software engineering. The dynamics of relational data over time can also be of interest, requiring time-varying graphs.

Creating effective visualizations for these diverse features is a challenge that has generated a lot of ideas from researchers and developers to address them. Whilst case studies and computational experimentation go some way to validating these ideas, it is ultimately the usability of visualization that must determine their feasibility. There is then a strong demand for human performance evaluation that tests ideas with end-users on relevant graph visualization tasks.

The literature on user studies for graph visualization has naturally become large and diverse in response to the need for evaluation (see Figure 1). A review of the literature is now timely to make the top-level results more accessible. In this state-of-the-art report, we survey user studies that focus on graph interpretation, graph memorability, and graph creation.

The associate editor coordinating the review of this manuscript and approving it for publication was Xi Peng.

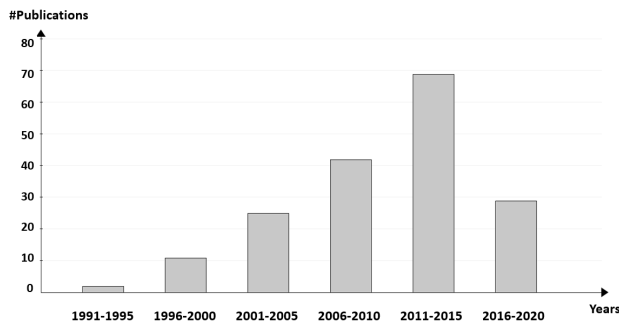


FIGURE 1. There is an increase of the number of publications related to empirical user evaluations in graph visualizations.

We present a top-level summation of results to provide a broad overview of common research themes and findings.

Yoghourdjian *et al.* [3] conducted a literature survey of user studies involving graph visualization. In comparison to the present review, the focus of their survey was on the size and complexity of graphs in user studies. It sought to answer the question of ‘what is an appropriate complexity of graph for user studies?’ In contrast, the present paper provides an overview and broad review of literature on user studies in graph visualization.

In practice, much of the work reviewed presents research that evaluates concepts used in graph visualization by means of human performance and judgment. As such, the present review provides insight into which techniques and concepts have proven effective, and which have not.

We structure this article by first describing background information and useful terminology used throughout this article (Section II). A data model on graphs and their properties is given in Section III. We then describe the scope of the review as well as the search and categorization of literature in Section IV. The review of literature is then presented in Section V, followed by general discussions as well as future research challenges in Section VI. Section VII concludes the paper.

II. BACKGROUND AND TERMINOLOGY

Graph data comes in a variety of forms such as planar or non-planar, multi-, bipartite, clustered, compound, or dynamic graphs. Also, hierarchies are considered a special type of graph. The edges can have properties for different graph classes, for example, directed or undirected edges, multiple edges between the same two vertices, or hyperedges. Vertices and edges can carry additional attributes, which may also have an inherent time-dependent nature. Moreover, the topology of the graph can be important and additional properties such as graph density (local or global density) can play a crucial role when looking for ways to analyze the graph data – either algorithmically or visually.

Many attempts have been made to visually encode such graph data [1], [2]. Depending on the inherent properties in the data and the task to be performed, some visualizations

may be more suitable than others. For example, a matrix visualization is considered better for dense graphs and cluster detection, whereas node-link diagrams are better for sparse graphs and path-related tasks. The problem in node-link diagrams is the increased visual clutter produced by the many links and link crossings if the layout is not well chosen. If a layout or vertex ordering is not done in the right way, visual clutter or hairball-like structures may be the result in node-link diagrams as well as unstructured adjacency matrices.

When a graph is visualized, the visual metaphors for the vertices, edges, time, and additional attributes are not the only important considerations. The medium on which graphs are displayed as well as the means of user interaction can also have an effect. Moreover, the task [4] to be performed by the user can change which visualization technique is most effective. Tasks are equally important for graph interpretation, graph memorization, and graph creation.

The representation of graph data started early with the work by Euler [5] and has since become a research discipline, with the International Symposium on Graph Drawing being held for the 28th time in 2020. Although this event typically covers approaches for node-link diagrams, their algorithmic runtime complexities, as well as algorithmic improvements, more and more empirical user studies are conducted to also investigate the task performance of end-users and how well graph diagrams are perceived. Graph drawing or graph visualization along with their various applications reach into several fields, including information visualization, visual analytics, diagrams, and human-computer interaction. Consequently, empirical user evaluations of graph visualizations can be found in nearly any journal, conference, or workshop that deals with the visual encoding of relational data. Therefore, writing this state-of-the-art report required an extensive literature search. We summarize the outcomes and take-aways from these studies, whereas Lam *et al.* [6] and Isenberg *et al.* [7] look more into several scenarios around studies in information visualization. Moreover, there are some general surveys on evaluating graph embeddings [8].

Throughout this article, we use the following terminology:

- **Hypotheses:** The term ‘hypothesis’ refers to a certain kind of assumption as a concrete description of what will be expected to happen in a user study.
- **Study methodology:** The term ‘study methodology’ means an evaluative methodology as described by Carpendale [9] such as controlled laboratory, uncontrolled, and longitudinal experiments.
- **Data:** When we use the term ‘data’, we refer to the graph data that is to be visualized, while this visual depiction has to be interpreted, to be memorized, or to be created. We distinguish ‘artificially generated data’ from ‘real data’.
- **Tasks:** When using the term ‘task’, we describe the user task that has to be solved when simultaneously interpreting, memorizing, creating, or interacting with a graph visualization [4].

- **Variables:** The term ‘variable’ expresses either ‘independent’, ‘dependent’, or ‘confounding variables’.
- **Participants:** The term ‘participant’ means the person who is taking part in a study who can be a layman or a domain expert, male or female, young or old, and so on.
- **Instrumentation:** The term ‘instrumentation’ describes which technologies are used to conduct the study such as the display media or data measurement devices (traditional vs. eye tracking).
- **Analysis:** By ‘analysis’ we express how the data is statistically evaluated such as statistical tests on dependent variables, but also visual analysis can be important such as statistical graphics like bar charts, box plots, or heatmaps and gaze plots for eye movement data.

III. DATA MODEL AND GRAPH PROPERTIES

We mathematically describe and model which kind of graph data can be visualized and, hence, can build the basis for an empirical user study in graph visualization (Section III-A). The graph data can have additional properties and each graph belongs to a certain graph class (Section III-B). If many stimuli of similar characteristics are required in a user study, the study designer has to apply useful data generation models (Section III-C). Moreover, the mappings of the graph data to a certain layout following aesthetics graph drawing criteria are frequently explored in user studies, which also requires mathematical modeling of these terms (Section III-D).

A. GRAPH DATA

When evaluating graph visualizations, we first have to understand which kind of data we are exploring before visually mapping the characteristics of it to graphical features. In the scope of this work, we model a static graph mathematically as $G := (V, E)$, which is a pair consisting of vertices V and the relationships $E \subseteq V \times V$ among them, which are denoted as edges. If more than one edge can exist between two vertices, we call this graph a multi-graph. If a single edge can connect more than two vertices at once, this edge is denoted by the term hyperedge.

A graph can also be extended by looking at the properties of the edges, which can be either directed or undirected. Moreover, edges can be accompanied by a quantitative attribute to which we refer as the weight of an edge. If we deal with weighted and directed edges, the graph is oftentimes called a network in the literature. If a list of attributes is attached to either vertices or edges or both, we speak of a multivariate graph. Mixed graphs may also contain an edge set E that allows a mixed edge type, i.e., directed, undirected, weighted, multi-, and multivariate edges.

In the context of this work, we also need to model a dynamic graph defined as a sequence $\Gamma := (G_1, G_2, \dots, G_m)$ of $m \in \mathbb{N}$ static graphs. Each G_i consists of a pair of vertices and edges, i.e., $G_i := (V_i, E_i)$, where each static graph can have one of the properties described above.

B. GRAPH CLASSES AND SPECIAL PROPERTIES

A graph cannot only be classified by the properties of its vertices and edges, but also by its topology and other special properties. A graph can belong to the class of planar graphs, allowing us to draw a node-link diagram of it in the two-dimensional plane without link crossings. The graph might be bipartite, i.e., the vertex set can be subdivided into two disjoint subsets where edges only exist between vertices from different groups. Consequently, edges within a vertex group are not allowed. A directed graph may be free of cycles, i.e., an acyclic graph. If an undirected graph is acyclic, we call it a hierarchy where one vertex is the designated root vertex of this hierarchy. Moreover, a compound graph has an additional hierarchical organization among its vertices, i.e., there are two different sets of edges: adjacency edges (in the graph) and inclusion edges (from the hierarchy of graph vertices).

We can also describe a graph by the ratio of edges per vertex. For example, if the number of edges grows in a squared manner compared to the number of vertices, we call it a dense graph, otherwise a sparse graph. If there is a group of vertices that has a dense behavior, we call it a cluster. If in this group, all edges between all vertices are present, we denote this by a clique, or n -clique if the number of vertices is n . For more descriptions of graph properties, we refer to Battista *et al.* [10] and Kaufmann and Wagner [11].

C. DATA GENERATION MODELS

Data generation models are required to guarantee similar characteristics for the graph stimuli in a study. In particular, when many graph datasets have to be shown (e.g., in a crowdsourcing experiment), the study designer cannot always generate sufficient graph data by hand nor source appropriate real-world datasets.

Graph data generation was first described by Erdős and Renyi [12] with the concept of ‘random graphs’. Each of the $n^2 - n$ possible edges in a graph is given a certain probability p . Although this model is simple, it is not applicable to real-world graph datasets. Using the model of Ware and Bobrow [13], graphs are generated by randomly adding edges to either one or two vertices for each other vertex. Single edges occur with a probability of p percent, whereas the others occur with a probability of $100 - p$ percent.

If a power-law distribution of the node connections to other nodes is required, i.e., a scale-free property of the graphs, another generation model has to be selected such as the Barabási-Albert model [14] or the Watts-Strogatz model [15].

D. VISUAL STIMULI

When a graph is represented in a visual form, its components, i.e., vertices and edges, have to be mapped to an output medium, i.e., to a visual display or by using sonification (for visually impaired people). This aspect demands a general mapping from data to visualization. In the context of graph visualization, we refer to both simple graph depictions (e.g., a graph drawing or a matrix representation) and more

complex graph presentation systems (i.e., systems that permit interaction, exploration, etc.).

In particular, when node-link diagrams are used, the graph visualization typically needs a layout algorithm. The chosen layout can affect the aesthetics of the visualization, the effectiveness of task performance, or both. For matrix visualizations, a layout algorithm is not needed, but rather a good vertex ordering that, for example, allows one to find clusters among a group of vertices.

We model a graph layout L as a function that maps graph vertices $v \in V$ to certain positions of the display space. The depictions of the edges between the vertices can be chosen from a certain repertoire of edge representation styles [16]. For the layout of the vertices there are several options like radial, circular, hierarchical [17], or force-directed ones [18] to mention some of a longer list. Random layouts are also sometimes used in a graph user study to have a basis for comparisons to sophisticated layout algorithms.

IV. SCOPE AND METHODOLOGY

We first describe the scope of this work (Section IV-A), then how and where we searched for appropriate literature (Section IV-B). The analysis and categorization of the resultant list of papers are then described (Section IV-C) and, finally, filtered for relevant work (Section IV-D).

A. SCOPE OF THIS WORK

In this article, we survey work on empirical evaluation of graph visualizations in which the user is involved. However, we do not include papers in this survey that only evaluate the presented graph visualization based on case studies without measuring user performance.

Moreover, this survey takes into account only papers that cover one or more topics related to the described graph data characteristics in Section III, including vertices, edges with directions, or the topology of a graph, i.e., if it is rather sparse or dense. Additionally, an inherent time dimension might be challenging to visually represent and, consequently, maybe also difficult to empirically evaluate. Many visualization problems can be transformed into a graph visualization problem, but we only add those papers that directly deal with a graph visualization problem. For example, in a broad sense, visualizing multivariate data might be understood as a graph visualization problem, if we interpreted the inherent tabular property of the multivariate data as an adjacency matrix and, consequently, as an underlying graph structure.

B. LITERATURE SEARCH AND COLLECTION

We checked the proceedings and issues of all main journals, conferences, books, book chapters, and workshops, some of them listed in the following:

• Journals

- Journal of Graph Algorithms and Applications
- Information Visualization
- IEEE Transactions on Visualization and Computer Graphics
- Computer Graphics Forum

• Conferences

- Symposium on Graph Drawing
- Joint Eurographics-IEEE VGTC Symposium on Visualization (EuroVis) [1999-2004: VisSym; since 2008 a special issue of Computer Graphics Forum]
- International Conference on Information Visualisation
- IEEE Symposium on Information Visualization (InfoVis) [since 2006 a special issue of IEEE Transactions on Visualization and Computer Graphics]
- IEEE Pacific Visualization Symposium (PacificVis) [2001-2004: InVis.au; 2005-2007: APVIS]

All articles and papers were read and manually scanned for their relevance for graph evaluation, which we recorded in a database. Additionally, we manually tagged all relevant papers, which finally led to our proposed classification. From each paper, we followed the citations and the references included in them by using the Google Scholar service. This procedure quickly leads to a fairly complete list of existing research in the field of empirical user evaluation of graph visualization.

C. ANALYSIS AND CATEGORIZATION

For the tagging process, we followed a similar approach as in the state-of-the-art report by Beck *et al.* [1] on dynamic graph visualization. We first used tags instead of categories, which allowed us to assign a list of tags instead of having to choose a single category. Before finding a suitable classification, we structured the defined tags into categories.

After we found a large portion of the existing relevant papers, i.e., about 110 papers, we discussed internally the found tags and, finally, came up with the high-level classification of graph interpretation, graph memorization, and graph expression and creation. For the subcategories, we did a similar process but on a more fine-granular hierarchy level. The remaining papers were then tagged one-by-one and added to one of the categories.

D. FILTERED LIST OF LITERATURE

At the date of submission of this article, our database contained 239 papers on empirical user evaluation in graph visualization, covering work from 1995 until 2020. From a formerly larger list of papers, we removed those that did not fulfill all criteria required for getting included in this survey. Those were papers that were off-topic such as computational evaluation papers, that only contained a case study on graph visualization, or that did not deal with graph visualization at all. Those papers might have been incidentally added to our database because a first quick manual scan through the paper was not accurate enough for us to identify it as a user study paper.

The papers finally added to the database were read by one of the authors, tagged, and categorized. The following section presents the result of this literature search and classification process.

V. CLASSIFICATION OF EMPIRICAL USER EVALUATION IN GRAPH VISUALIZATION

Our top-level classification is based on the high-level tasks addressed in the experiments, typically: interpretation, memorability, and expression (i.e., creation). Thus, our focus is on the effect of the visualization on the user (rather than on the graph visualization itself) and, thus, on the usefulness of visualizations, and on experimental outcomes that can inform the design principles used for enhancing or creating useful graph visualization techniques.

By interpretation, we mean the ability to understand the graph drawing and/or the information incorporated there – typical research questions in this category might be “Does the presence of edge crossings hinder the ability to find shortest path routes between nodes?” or “Does the presence of edge crossings hinder the interpretation of the friendship relations in a social network?”

By memorability, we mean the ability to remember a graph drawing and/or the information encoded therein – typical research questions in this category might be “Is a dynamic graph consisting of a small number of graphs easier to remember than one that contains various graphs?” or “Is it easier to remember who holds the power in a social network when it has been presented using a force-directed algorithm rather than a circular layout?”

By expression, we mean the ability to create a graph drawing that represents information. Studies in this category are usually not so concerned with the ability to create the drawing, but the manner in which it has been drawn, so typical research questions might be “Do users conform to common graph layout principles when creating their own drawings?” or “How do users represent close friendship networks in a social network?”

Most studies we identified focused on one of these three high-level tasks. In some cases, those tasks relate mainly to the structural form of the graph (e.g., shortest paths, cut nodes, nearest neighbors), and in others they relate to the relational information included therein (e.g., class inheritance, friendship cliques, data flow).

A. GRAPH INTERPRETATION

Once a graph is visualized, i.e., represented to the user, an important question is whether the graph as it is displayed is readable, understandable, and effective. To find this out, several visual features of this graph have to be tested for their suitability. This is typically done by varying some parameters in the display, e.g., the visual appearance of vertices and edges, the graph layout, or the visual mapping of an inherent time dimension if one is dealing with dynamic graphs. All research papers focusing on user experiments taking these criteria into account, are categorized under the aspect of ‘graph interpretation’.

1) VISUAL DESIGN

Since any graph consists of vertices and edges, the simplest kind of visualization is one that visually encodes vertices,

edges, and, thus, the structure of the graph. We first discuss vertex and edge representation styles and how they were evaluated (Section V-A1.a). The next step is to survey work that takes into account the visualization of a complete graph, i.e., we describe evaluations of basic visual metaphors such as node-link diagrams, adjacency matrices, or adjacency lists (Section V-A1.b). As an additional dimension, an ordered list of graphs can be visualized, having a pre-defined temporal or sequential order (Section V-A1.c). Accordingly, this section describes work on the visual design of vertices and edges for single and multiple graphs.

a: VISUAL PROPERTIES OF VERTICES AND EDGES

When relational data is visualized, the designer must be able to graphically express relationships between objects. For node-link diagrams, this is done by representing graph vertices as visual nodes (squares, circles, triangles, and so on) and edges in between as explicit links (lines of a certain thickness and shape). For adjacency matrix visualizations, a relation is typically visually encoded by placing a color-coded cell at the row and column intersection point that indicates the corresponding related vertices. Such an implicit relation encoding is also used for adjacency lists in a similar way, expressing the number of adjacent edges.

In many scenarios, node-link representations are used, requiring an explicit link encoding of the presented relationships. A general problem when using links to represent relations between objects is the increasing amount of visual clutter [19], in particular, when graphs become dense and/or the applied layout algorithm cannot effectively handle the corresponding graph structure.

The impact of edge representations on user performances has already been studied in empirical evaluations. For example, the shape of links presenting edges might be varied to understand if, for example, straight-line drawings or curved links perform better. Bar and Neta [20] investigated in a user experiment if people rather prefer curved visual objects. By asking 14 participants, they came to the conclusion that whether a contour is sharp-angled or curved has a critical influence on people’s attitude to the presented stimulus. This is not a graph visualization study, but it might be also of interest if the same assumption holds for edge representations.

Pohl et al. [21] conducted an eye tracking experiment with 36 students to find out if an orthogonal [22], a force-directed [18], or a hierarchical layout [17] performs better. Ninety-degree link bends were used for the orthogonal drawings, straight links in the force-directed diagrams, and curved links in the hierarchical graph representations. Node search, link, subgraph, and 4-clique existence, and node property tasks were asked. The main result of this study was that the force-directed layout with straight links outperformed the orthogonal and hierarchical layouts with ninety-degree link bends and curved links for all five tasks.

Holten and Van Wijk [24] studied standard links with arrow-heads, light-to-dark, dark-to-light, green-to-red, curved, and tapered edge representations (see Figure 2 for

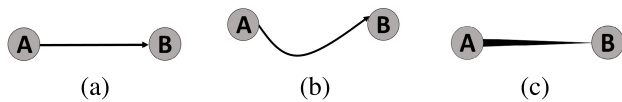


FIGURE 2. Different edge representation styles as evaluated in the work by Holten *et al.* [23]: (a) Standard straight line with arrow-head. (b) Curved representation. (c) Tapered edge representation style.

three edge representation styles from the study). Single-cue directed edge representations with 30 participants and a follow-up multi-cue experiment with 15 participants were conducted. Single-step and two-step connections had to be determined. The standard arrow-heads representation performed worse than the others, whereas the tapered edges performed well. Curved links were the worst representation based on completion times and error rates. They could not find a clear performance benefit with the multi-cue directed edge visualizations. In a follow-up study, Holten *et al.* [23] also investigated the effect of animated links and biased-curvature by testing 27 participants. Also here, the tapered and animation style performed better than biased-curvature ones. Another follow-up study by Holten *et al.* [16] also tested the performance of textured links, i.e., glyph representations, compared to tapered and animated ones with 25 participants. The results also showed that the glyph representation could not keep up with tapered and animated approaches.

Okoe and Jianu [25] conducted two studies to demonstrate the effectiveness of their crowdsourcing platform and evaluation framework for graph user studies. One of them partially replicated the study by Holten and Van Wijk [24] in which three styles of directed edges were used: tapered, arrow-head, and curved. With a sample of 62 participants, the finding that curved edges were less effective was replicated. Unlike the original study, however, arrow-heads were found to be more accurate than tapered edges, but tapered edges were faster for the one and two-step path search tasks.

Xu *et al.* [26] studied the impact of curved links on readability tasks. They compared straight links, links with different curvature levels, and mixed straight and curved links. In two separate experiments, they first investigated links with the same curvature level for all links present in a graph with 28 participants, whereas in the second, 65 participants were asked to answer tasks in node-link diagrams in a Lombardi-style force-directed representation. A uniform curvature level had a negative impact on graph readability, which increased with additional curvature. In contrast, the Lombardi effect had no significant impact on user performance when compared with straight links.

Couch [27] tested curved and straight links in force-directed graphs by recruiting 32 participants. Node distances in node-link layouts with straight and curved links had to be judged. In the end, participants were asked to rate which graphs were easier to read, more pleasant to view, and faster to explore. As a result, 30 of 32 participants found straight link graphs faster to explore, while 23 out of 32 found straight graphs easier to read. Only on the aesthetics judgment,

i.e., which graph was more pleasant to view, 16 out of 32 were in favor of the curved links. Consequently, also this study shows that curved links might look aesthetically appealing, but from a graph readability perspective, those perform worse compared to straight links.

Huang *et al.* [28] also studied the effects of edge layout styles on performance and user preferences. Curved and straight edges, as well as layouts with and without edge crossings, were compared. A sample of 26 participants showed that users performed fastest when edge-crossing layouts were used for graph interpretation tasks. No reliable differences were found for accuracy or perceived effort. User preferences for both performance and aesthetics were better for layouts with no edge crossings, and for curved edges when edge crossings were present.

A totally different kind of edge representation style was introduced by partial link drawings. Rusu *et al.* [29] used breaks at link crossings. Those gaps can then be completed in the mind by following the Gestalt law of 'closure'. In a user study with 14 participants, all nodes connected to a highlighted node had to be identified with and without breaks at possible link crossings. Qualitative ratings for this task were asked from the participants, indicating no differences.

Burch *et al.* [30] tested straight partial links, but in traditional and tapered edge representation styles. Forty-two participants answered path-related and link adjacency property tasks in force-directed layouts with varying sizes. Partial link drawings can lead to shorter task completion times but also to higher error rates. Seventy-five percent link length was uncovered as being a good percentage for both completion time and error rate for the tested tasks and graph sizes. Burch [31] extended these findings for radial graphs. In an experiment with 53 participants, the accuracy of a link-following task was investigated in which the length and direction (i.e., angle) of the edge were varied. It was found that accurate judgment was possible for edges that were 65 percent or more of the total possible length. Edges that were closest to horizontal or vertical orientations were also most accurately judged. Edge length was found to be less important for accurate judgment for horizontal and vertical orientations.

A similar study using partial edges was conducted by Sathiyarayanan and Pirozzi [32]. They compared Euler diagrams of networks that involved either complete or partial edges. A sample of 20 participants performed standard interpretation tasks on the graphs. It was found that partial edges produced significantly fewer errors (at $p < .05$ standard level; the paper reports at $p < .005$ level) but were equivalent to complete edges in terms of response time. Eighty percent of the participants also indicated a preference for the partial edges.

Summary: There seems to be a tendency to use straight links [21], [24], [27] and those visualized in a tapered representation style [23], [24]. However, arrow-heads instead of tapered links might be more accurate [25]. Similar findings hold for partial links [30]–[32], however, they should

be between 65 to 75 percent of length to be faster and more accurate than complete straight links. On the negative side, curved [25]–[27] and animated links [23] as well as glyph-based approaches [16] seem to be worse than straight links, but from an aesthetics perspective, they have some benefits [27], [28].

b: METAPHORICAL GRAPH REPRESENTATIONS

Depending on the topological properties of a graph, there are several general options for graph representation: using a node-link diagram, an adjacency matrix, or an adjacency list (see Figure 3 for a node-link diagram and a corresponding adjacency matrix). Also, hybrid approaches may be useful to reduce visual clutter caused by link crossings, but still have a fairly space-efficient representation such as MatLink [33] or NodeTrix [34].

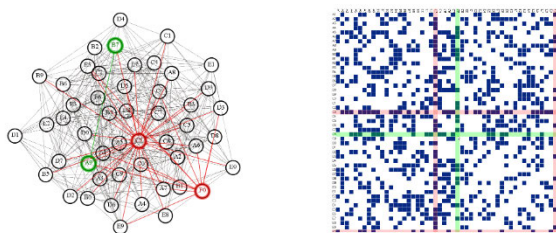


FIGURE 3. Evaluation of node-link diagrams and adjacency matrix visualizations. Image reprinted from Ghoniem *et al.* [35], © 2004 IEEE.

Node-link diagrams are by far the most prominent visualization style for relational data. The prevalence of node-link diagrams might be because of early work by Euler [5], who found an abstraction for the problem of ‘The Seven Bridges of Königsberg’, which he modeled as a node-link diagram.

Many years later, adjacency matrices were introduced [36], which are typically used when graph data becomes dense, i.e., many edges exist. Adjacency lists [37] are only rarely used for graph visualization since they rather serve as a space-efficient internal data structure to handle graph data.

Ghoniem *et al.* [35], [38] evaluated if node-link diagrams or matrix visualizations (see Figure 3) perform better for typical graph readability tasks by recruiting 36 participants. For graphs larger than 20 vertices, adjacency matrices performed better than comparable node-link diagrams. Generic tasks in this study involved typical estimation tasks on node and link number, or the most connected node. Also, label search tasks for nodes and links were checked as well as finding common neighbors or typical path-related tasks between start and target nodes. Only for the path-related tasks, node-link diagrams were the favorite visual metaphor for graph data.

Okoe and Jianu [25] performed a partial replication of the study by Ghoniem *et al.* [35] with their crowdsourcing platform. Node-link and matrix representations were compared for a neighborhood search task. With a sample of 112 participants, it was found that the matrix representation produced more accurate and faster responses. A follow-up

study by Okoe *et al.* [39] involved more tasks and participants. With data from 835 people, they found that node-link graphs generally outperformed matrix representations on path-related tasks, but matrices were best on common neighbor and group tasks.

Keller *et al.* [40] studied connectivity models by evaluating matrices and node-link diagrams in two experiments. For the first one, 21 participants were recruited and it was checked if matrices were readable and if the size, the density, and the directionality had any influence on response times and error rates. An online experiment was conducted in which the participants had to count incoming or outgoing links. All factors had a significant impact on the performance measures, i.e., the dependent variables. For the large and dense graphs, the completion times for solving the task were much higher. In a second experiment with 16 participants, matrices were compared to node-link diagrams. Tasks involved node and link selection, counting the number of incoming and outgoing links as well as common neighbors, and finding the length of the shortest path between highlighted nodes. The path reading task could be solved faster in node-link diagrams for small graphs, a result very similar to those by Ghoniem *et al.* [35], [38].

Henry and Fekete [33] designed MatLink, a hybrid graph visualization, combining adjacency matrices overlaid with node-link diagrams using curvature for the links. The question was if hybrids can keep up with the individual metaphors, i.e., node-link diagrams and adjacency matrices. Thirty-six participants had to answer social network-related tasks. MatLink performed significantly better for most of the tasks, in particular, for path-related tasks. Also, adjacency matrices performed worse than node-link diagrams for path-related tasks.

Brain connectivity analysis can be visually supported by either node-link or matrix representations, which was studied by Alper *et al.* [41]. Eleven participants had to answer typical tasks such as edge weight changes, connectivity changes, and regions subject to the most changes. Those tasks had to be answered in both modified node-link diagrams and modified adjacency matrices that visually encode two graphs at the same time by color coding. All tasks were answered more accurately and faster for the adjacency matrix stimuli.

McBride and Caldara [42] tested tables (a text-based representation of adjacency matrices) against graphs in a radial layout. Eighty-six participants were asked to select a single criminal to arrest in the displayed network. The results showed that node-link diagrams are the better choice for this task due to the much lower completion times.

Hlawatsch *et al.* [37] also tested the suitability of list representations for displaying dynamic graphs. In a user experiment, they compared node-link diagrams, adjacency matrices, and adjacency lists. The existence of a link, the equal distribution of incoming and outgoing links, and a weight-related task were asked on static graphs. Another weight-related task had to be answered in a dynamic graph. Twenty-four participants were involved and performed quite

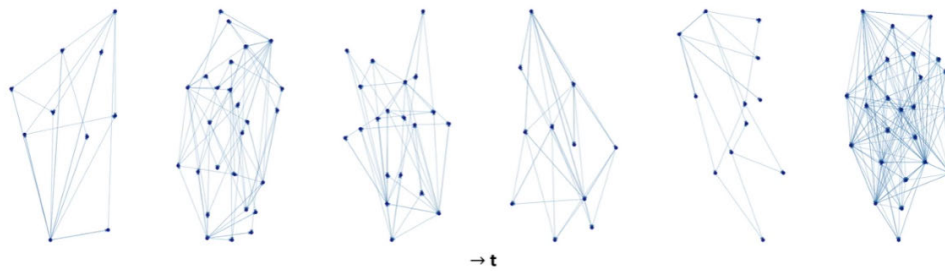


FIGURE 4. A time-varying graph consisting of 6 time steps visually depicted in a node-link visual metaphor. Investigating the effect of the visual variables in a dynamic graph visualization is a challenging aspect in a user evaluation. The time dimension adds another independent variable to a user study worth exploring.

well with adjacency lists compared to the other visual metaphors. However, for the link existence task, completion times for adjacency lists were high compared to node-link diagrams and adjacency matrices. For the weight-related tasks, the adjacency lists led to lower completion times, for both the static and dynamic graphs. The error rates showed a similar behavior.

Summary: There seems to be a tendency to use adjacency matrices [25], [35], [38], [41], in particular, if the datasets are growing larger and the graphs become denser. However, node-link diagrams have some benefits for path-related tasks [33], [35], [38]–[40], [42], but only if the graphs are not too large. Adjacency lists [37] only seem to be suitable visual concepts for a few very specific tasks.

c: VISUAL ENCODINGS OF A GRAPH SEQUENCE

Not only a single graph can be of interest for a visual design, but also a multitude of them (number of graphs ≥ 2), typically occurring as an ordered list or sequence of graphs. This additional dimension is a challenging problem for graph visualization (see Figure 4). A sequence of graphs can just have an inherent order, but it can also be time-based, making it a dynamic graph or a time-varying graph.

Beck *et al.* [1] surveyed existing work on dynamic graph visualization. Although there are various approaches for visualizing time-dependent relations, Beck *et al.* found two major strategies, namely time-to-time mappings and time-to-space mappings. A big issue for dynamic graph visualization is the preservation of a viewer's mental map when one tries to inspect dynamic graph visualizations. Consequently, most of the user studies in this field focus on trying to find out which representation is suited best for the viewer with respect to mental map preservation. Moreover, several aesthetic criteria and application requirements for dynamic graph visualization should be followed [43], [44], sometimes acting in a trade-off behavior to the mental map preservation concept.

Bridgeman and Tamassia [45]–[47] describe a user study in which they explored how to compare graph drawings. Similarity measures were proposed and validated in a user experiment. By investigating agreement between the metric

and human judgment in the user experiment, they evaluated how humans perceived similarity, which was then compared to the formal similarity measures. 103 students were given three tasks, split into a rotation part, an ordering part, and a difference part, in which orthogonal node-link diagrams had to be visually compared. It was found that point positions are important to the perception of similarity, but less significant for ordering.

The performance of difference maps in dynamic graphs was studied by Archambault *et al.* [48]. Animated, slide show, and matrix small multiples graphs with and without difference maps were used. Twenty-five participants were asked about the local topology-based evolution of node degrees, edge appearances, global edge trends, and global topology-related tasks such as path tracking over time. Difference maps produced significantly fewer errors when judging the number of inserted or removed edges as the graph evolves over time. The participants preferred the difference maps in all tasks.

Zaman *et al.* [49] recruited 16 participants for two user studies on a hierarchically laid out graph sequence using animation, dual view, difference layers, and a relative re-layout. The first study tested a node insertion task and showed that difference layers are best and the dual view is worst for error rates and completion times. In a second study, the participants had to detect shifted nodes in two versions in a single-view animation, a dual view with a difference layer, and by using a combination of the dual view with a difference layer combined with animation. In this study, the difference layer was the worst technique while the other two were similar.

The visualization of graphs with associated time-series data was studied by Saraiya *et al.* [50]. In their study, 40 participants were shown a single static graph, two graphs in a sequence, and a longer sequence of graphs. The explicit tasks asked, for example, for node values in a static graph, node value changes in two graphs to be compared, topology trends, search over time, or outliers in a dynamic graph consisting of more than two timesteps. The results showed that overlaying data on the graph nodes for each timepoint performed more accurately for single timepoint analyses and when two graphs

were compared. A simultaneous overlay with data for many timesteps can lead to more accurate and faster performance for outlier searches among the vertices. For topological tasks, single views are better than multiple views.

Purchase *et al.* [51] explored the importance of the mental map for dynamic graph visualization. Twenty students were asked to solve link adding, node changing, and timestep finding tasks. An animated hierarchical dynamic node-link layout was used in which the delta conditions served as independent variables. Low delta values, which better maintain the mental model, produced better user performance for the tasks. In another work, Purchase and Samra [52] found out that extremes are better for dynamic graph visualization, but it also depends on the individual preference.

Two dynamic graph layout algorithms were studied by Saffrey and Purchase [53]. The mental map builds the key aspect for each algorithm. This study also tested different mental map conditions ranging from high, medium, and low to zero. Node and timestep search tasks were answered by 21 participants. No significant results were found for the completion times, but for the error rates, the high mental map condition produced the most errors, whereas the zero mental map condition produced the least.

The mental map condition for animation and small multiples was investigated by Archambault *et al.* [54]. Local and global properties during graph evolution were tested in this study. The tasks were performed faster with the small multiples representation than with the graph animation for all tasks. For error rates, animation was significantly better when asking if nodes and edges were added to the same time slice. The preservation of the mental map had only little influence on the performance measures for both animation and small multiples. Archambault and Purchase [55] surveyed experiments and findings on the mental map preservation. They gave recommendations in which case and for which tasks the mental map supports a human user. Further challenges in this field of research were discussed in their work [56].

The ability of users to track paths using animation or small multiples was involved in a subsequent study by Archambault and Purchase [57]. With a sample of 28 participants, no reliable differences were found between small multiples and animation in terms of speed or accuracy. There was, however, a nominal trend for animation performing better, especially for the condition without mental map preservation. Responses were reliably faster and more accurate when the mental map was preserved. A follow-up study [58] focusing only on the condition with no mental map preservation, again with 28 participants, found that animation produced more accurate and faster responses than small multiples on the path tracing task.

Small multiples were further investigated by Archambault and Purchase [59] for cascading of node attributes in directed dynamic graphs. Hierarchical and force-directed layouts were compared, as were animated (i.e., time slider) and time slice presentations of graph evolution. With a sample of 21 participants, it was found that

time slices led to faster performance of a cascade graph interpretation task. It was also found that the hierarchical layout led to faster and more accurate performance for animated presentation, as well as faster performance for time slice presentation. The presentation of cascade history was also investigated but did not affect participant performance.

Boyandin *et al.* [60] conducted a qualitative user study with 16 participants to explore animation and small multiples representations for temporally changing flow maps, which are in some way related to graphs. With animation, the participants were able to detect findings more locally, whereas with the small multiples setting, they also found insights in longer time periods.

Key and Diehl [61] studied interactive dynamic graphs by controlling the presentation speed, labels, and tooltips. These factors may have an influence on the performance of user comprehension of an evolving graph. A sample of 111 students inspected animated node-link diagrams and answered twelve multiple-choice comprehension questions. The adjustment of the presentation speed was rarely used by them while displayed labels performed better compared to the tooltip option.

Shi *et al.* [62] introduced a 1.5D egocentric dynamic network visualization that they evaluated in a user study by comparing it to a small multiples approach and animation. Twelve participants performed tasks testing topological network and temporal features. The 1.5D approach performed well for completion times and error rates, whereas animation was the slowest technique.

Kondo *et al.* [63] introduced Glidgets, an interactive glyph-based visualization for dynamic graphs that they compared to a traditional time slider technique. The 8 participants had to find the timestep where a certain node had a special property. The glyph-based technique did not outperform the time slider technique.

Bach *et al.* [64] studied typical ‘where’, ‘what’, and ‘when’ questions for the GraphDiaries technique, video animation, and a flipbook technique. There was no significant difference considering error rates for the trend detection task, which led to a follow-up experiment with the same 18 participants. The task completion time increased for GraphDiaries, while the error rates were significantly reduced and performed better than video animation.

Hybrid and non-hybrid techniques were compared by Rufiange and McGuffin [65]. In their DiffAni tool, they showed, by recruiting 12 students, that the hybrid representation has some advantages over non-hybrid techniques such as graph animation.

Apart from node-link approaches, matrix-based representations have also been designed for dynamic graph visualization. Burch *et al.* [66] compared a Cartesian [67] and a radial variant [68] of a visualization technique for displaying weighted directed dynamic compound graphs. For their eye tracking experiment, 35 students were recruited who had to answer typical graph-specific tasks such as correlation and counting questions. The Cartesian diagram outperformed the

radial one for most of the tasks. Only the correlation tasks were answered more accurately in the radial variant.

Summary: There seems to be a tendency to use static (time-to-space) mappings (like time slices, small multiples, 1.5D, and so on) for a dynamic graph [54], [59], [60], [62], in particular, if comparison tasks [50] have to be conducted to identify trends in the time-varying graphs. The mental map preservation [51]–[53], [55]–[57] is an important concept in this context, for example, by keeping similar positions [45]–[47] or only showing the differences [48], [49], which might be problematic for graph animations or time slider techniques [63]–[65]. However, for path tracing tasks, animation could have some positive effects [58], but animation speed does not seem to play a large role [61]. Also the representation of the time axis could make a difference for the dynamic graph visualization [66].

2) LAYOUT

User studies directly evaluating the layout of a graph, also in comparison to others, are discussed in Section V-A2.a. User studies can also investigate how a layout, following certain aesthetic criteria, influences human performance, which is discussed in Section V-A2.b. Special clutter reduction techniques are described in Section V-A2.c.

a: LAYOUT ALGORITHMS

Blythe *et al.* [69] studied the layout effect in social networks on social grouping and actor centrality detection. Eighty participants were recruited and they were shown three out of five different layouts of the same graph. Significant effects of the layout were found, but there was no best layout candidate. Later, McGrath *et al.* [70] described a user study investigating the impact of the layout on the completion time. Sixty-one participants (graduate students) were asked to find a particular node, to activate a group, and to assign nodes to that group. Again, three out of five different layouts were compared by asking the task of assigning nodes to groups. The results show that spatial clustering has a significant effect on viewers' perception of group existence in networks, i.e., structural graph features are held constant while Euclidean spatial factors have an influence on the users' perception [71]. Several years later, McGrath and Blythe [72] studied the effects of layout and motion on viewers' perceptions and performance of displayed networks. The motion feature had a positive influence on the 133 viewers' perception of change. No effect of hierarchical versus spatially central layouts on error rates was identified.

The usefulness of grouped network layouts was also studied by Chaturvedi *et al.* [73]. An existing algorithm using a squarified treemap to layout network groups in a grid, showing inter-network edges connecting their centers, was compared to layouts developed by the authors. Two alternative layouts were formulated that were designed to better maximize space usage and reduce edge crossings and occlusions. The use of each alternative was determined automatically by the size and number of the networks involved. The pilot study with 9 participants showed that users believed the

alternative layouts were substantially more useful for inter-network interpretation tasks.

Homophily is an attribute of network groups and clusters that was investigated by Meulemans and Schulz [74]. Homophily is the degree of intra-cluster connections compared to inter-cluster connections. A force-directed algorithm was compared to a modified force-directed algorithm designed to separate cluster nodes to the left or right, and a bipartite layout algorithm that strictly separated cluster nodes and used straight edges for inter-cluster connections and arced edges for intra-cluster connections. A web-based mixed design experiment using 90 participants' data showed that homophily perception was better in the bipartite layout in terms of estimated deviation for actual score and response time. However, node separation could not be attributed to the difference in homophily perception because the modified force-directed layout produced the worst performance. Compared to homophily perception, the shortest path search task was faster and more accurate for the force-directed algorithm.

Forty-six participants were shown several graph layouts by Dengler and Cowan [75]. Semantic conclusions were given, obtainable by inspecting the layout. No semantic attributes were attached to the graph nodes. The researchers asked the question if semantic attributions were consistent or random and to identify those consistent attributions if they exist. The participants all agreed to the semantic content of specific graph layouts. No difference was found between experienced programmers and people only rarely working with computers.

Purchase [76] studied eight different layout algorithms. With the exception of one algorithm, all others did not show any statistical difference in user performance. In particular, for UML collaboration diagrams (which are a special kind of graphs), Purchase *et al.* [77] empirically studied the performance of 35 subjects by testing five UML notations in two variations each. A pseudo-code specification had to be matched with UML diagrams. Subjects preferred the more concise notational variants, a result which was also confirmed in Purchase *et al.* [78] by asking which of two complete notations is easier to understand.

Hierarchical layout was investigated by Körner [79], recruiting 12 female students. Interpretative questions such as the comparison of graph nodes had to be answered. A two- and three-stage exploration model were uncovered, two search stages for the graph nodes and a reasoning stage in the end that combines the information about the found nodes. Some years later, Körner [80] applied eye tracking to explore the eye movement behavior and visual attention to hierarchical graphs when asking tasks about target node search, relation reasoning, or both in combination. Such hierarchical graphs were read in a sequentially applied strategy.

Hierarchical layout was also one of four layouts compared by Didimo *et al.* [81] for directed graphs (see Figure 5). Specifically, overloaded orthogonal, orthogonal, hierarchical graph layouts as well as matrix representations were compared in a user study involving 21 participants. Path, cycle,

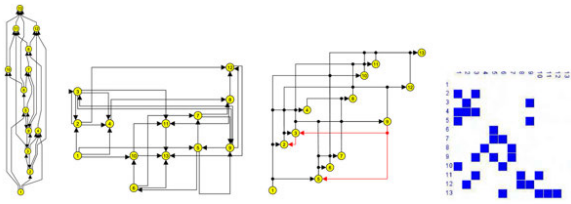


FIGURE 5. Evaluation of different node-link layouts and matrix visualizations. Image reprinted from Didimo *et al.* [81], © 2017 The Authors Computer Graphics Forum, © 2017 The Eurographics Association and John Wiley & Sons Ltd.

and edge finding tasks were performed to compare the efficacy of the layouts. Overloaded orthogonal layout produced the fewest errors, whereas matrix representation produced the most. For speed, the hierarchical layout was best followed by overloaded orthogonal, and matrix was again worst. All of the tasks followed this pattern except for a task to identify the degree of out-going edges in which the results were statistically equivocal.

Huang and Eades [82] used eye tracking to explore the effect of layout characteristics. A sample of 13 participants had to answer typical path-related tasks. Circle layout (all nodes on a single circle) and radial layout (all nodes on different radial layers) were tested for 12 graphs. The result of this study showed that the graph reading behavior is a very complex process that needs further investigation.

Different layouts for metabolic networks were investigated in a comparative study conducted by Bourqui *et al.* [83]. The performance of 22 participants for motif-search tasks was recorded in force-directed, hierarchical, and one specifically designed layout that takes existing metabolic representation conventions into account. The force-directed layout led to better user performances for the search task.

Sixty-nine participants were asked typical graph reading tasks in Spring embedder, Lombardi, and restricted Lombardi layouts by Purchase *et al.* [84]. In previous studies on edge representation styles, it was found out that curved edges do not perform well compared to straight line drawings. A similar finding holds for Lombardi drawings that are preferred by the study participants, but with respect to performance, Lombardi drawings perform poorly.

Dawson *et al.* [85] investigated the path tracing task for graphs in a force-directed layout. The authors conducted a user study asking 12 participants to complete 144 unique path tracing trials. By observing and characterizing human path-tracing behaviors, a predictive model of the search set for node-link diagrams was developed and validated.

Summary: There seems to be a tendency to use layouts with a spatial clustering for group identification tasks [70], [71], [73] or even motion to see layout changes [72]. Force-directed layouts [74], [83] seem to have benefits for group and motif-search tasks. Even semantics was useful to better explore a graph [75], for example, in the form of a UML diagram [76]–[78]. Strategic exploration patterns were detected in hierarchical layouts [79]–[81] and force-directed

layouts [85], also by using eye tracking, uncovering complex reading processes in radial graph diagrams [82]. Lombardi drawings are preferred [84], but due to their curved edges, they lead to bad performances. Some studies do not show significant effects for a certain layout [69].

b: AESTHETICS CRITERIA

A list of aesthetics graph drawing criteria was proposed by Bennett *et al.* [86], focusing on node-link diagrams. Beck *et al.* [43] introduced an aesthetics dimensions framework in which they also investigated aesthetics criteria for general visual metaphors such as adjacency matrices and adjacency lists.

In this section, we describe user studies that take into account the impact of aesthetics criteria on user performance. Although layout algorithms are already typically designed to follow a certain list of aesthetics criteria, we do not discuss here user studies that solely focus on evaluating different ‘standard’ layout algorithms, but we rather discuss research on explicit applications of these criteria on pre-computed graph layouts.

The validation of claims considering the optimization of layout aesthetics qualities was studied by Purchase *et al.* [87]. Various layout aesthetics were used to explore human understanding of graphs. As a result, Purchase *et al.* found out that an increase in the number of arc bends and arc crossings leads to a decrease in understandability.

Purchase *et al.* [88] recruited 49 students for testing different layout aesthetics for undirected graphs. Path-related tasks and graph interpretation tasks were asked on sparse and dense graphs. The influence of aesthetics including bends, crossings, and symmetry was evaluated, and it was found that both bends and crossings have a negative impact on task performance, while the effect of symmetry was not confirmed. In the same line of research, Purchase [89] asked which aesthetics has the greatest effect on user performance. To reach this goal, five aesthetics were tested and ordered by their relative importance. The minimization of link crossings was by far the most important criterion, followed by minimizing the number of bends and maximizing symmetry. There was a significant influence of maximizing minimal angles between links leaving a node and fixing nodes and edges to a grid. Purchase *et al.* [90] and Purchase [91] further experimented with such aesthetics-based graph layouts and found that some individual aesthetics influence human task performance, but when it comes to the performance of algorithms that were to optimize multiple aesthetics, it is inconclusive on whether one algorithm is better than another one.

A user preference study on individual aesthetics in the domain of UML diagrams [92] was conducted by Purchase *et al.* [93], [94]. The final ranking of the measured aesthetics showed that the ranking can be different from domain-independent studies. Thirty participants were recruited by Purchase *et al.* [95] in a comprehension study investigating UML class diagrams based on the most important drawing aesthetics. The semantics of the graph dataset

plays a crucial role and should be considered before generating a layout that follows a specific aesthetic criterion. In another article, Purchase *et al.* [96] describe how aesthetic criteria influence human performance by investigating individual aesthetics criteria, common automatic graph layout algorithms, and individual aesthetics criteria used for UML class and collaboration diagrams. Carrington *et al.* [97] followed this line of research on the layout and notation in UML diagrams. At least 30 participants were asked in several experiments on aesthetic and notational variations including bends, link crossings, width of layout, font type, text direction, orthogonality, inheritance metric, number of arcs, and directions. Fewer crosses and bends were significantly preferred.

Ware *et al.* [98] studied cognitive measurements of graph aesthetics. The Gestalt law of good continuity was found out to be of special interest, i.e., it should be guaranteed that longer paths through a node-link diagram should be kept as straight as possible (see Figure 6). Spring layout graphs were used to evaluate the task of finding shortest paths. Apart from continuity, link crossings were also found to influence user performance, in particular for long paths.

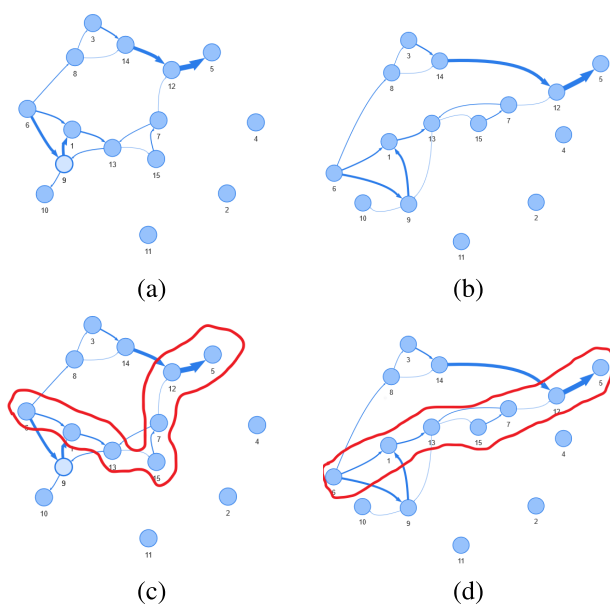


FIGURE 6. For path tracing tasks the longer paths through a node-link diagram (a,c) should be kept as straight as possible (b,d) taking into account the Gestalt law of good continuation.

Aesthetics criteria such as planarity, slopes, and levels in hierarchical graphs were studied by Körner and Albert [99]. Thirty participants were recruited for 4 different experiments answering questions on comparisons among the graph nodes and edges. The crossing of links was the most influential aspect for human performance and speed of comprehension.

The impact of link crossings and layout effects on sociogram perception was measured in a user study by Huang *et al.* [100], [101] and further explored in a questionnaire study [102]. Link crossings and drawing conventions

have a significant impact on group finding task performance. The questionnaire study showed that people tend to place nodes on top or in the center to indicate importance as an aesthetic criterion. Clustering nodes into groups indicates strong relations among nodes in that group.

Huang conducted an eye tracking study [103], [104] with 16 participants asking path searching and node locating tasks for 6 node-link graph drawings. Aesthetics criteria such as crossing angles and geodesic-path tendency were tested. Small angles can slow down and trigger extra eye movements, which lead to delays for path searching tasks. In contrast, crossings only have little impact on node locating tasks while the geodesic-path tendency shows that paths between two graph nodes are much harder to follow when there are many branches going toward the target node. In other experiments, Huang *et al.* [105] confirmed the existence of the geodesic-path tendency by applying eye tracking.

To investigate the effect of crossing angle, Huang *et al.* [106] conducted a controlled experiment in which 22 participants were recruited and asked to perform a path tracing task on graph drawings with varying crossing angles. It was found that the increase of crossing angles' size led to a decrease in completion time. This finding was further confirmed in another study with 37 participants on drawings of general graphs [107]. Huang *et al.* [108] summarized and illustrated the evaluation approaches that they used in these empirical studies, including questionnaires, eye tracking, and cognitive load studies, for evaluation of graph drawings.

Conflicting aesthetic criteria were the basis for a user study by Huang *et al.* [109]. Effectiveness of user performance can be improved when a layout algorithm takes into account more than one aesthetic criterion and finds suitable compromises. In their study, force-directed layouts (Spring embedder and BIGANGLE [110]) were compared, and 43 participants were asked to perform a path finding task. The BIGANGLE algorithm with multiple aesthetics in a compromise had significantly better user performance. In another study, Huang *et al.* [111] investigated the effect of angular resolution on human task performance and found that the smallest angle formed by any two neighboring edges is the best measurement for angular resolution.

An exploratory user study with 32 students on the effect of the relative importance between crossing number and crossing angle was conducted by Huang and Huang [112], [113]. In this study, not only task performance in the form of completion time and error rates was measured, but also cognitive load and visualization efficiency. The number of link crossings was found to be more important than the size of crossing angles.

Eye tracking was used in two experiments by Huang [114]. In the first experiment, 13 participants were asked to find separation levels between two families of a social network. The edge crossing effect had a negative influence on eye movements and user performance. In a second experiment, 16 participants were recruited to find a shortest path in node-link diagrams. An analysis of the video data from the

first experiment revealed that the decreased performance was not caused by the link crossings, but rather by whether the searched path is close to the geodesic path tendency.

In an effort to propose and validate an overall quality measurement, Huang *et al.* [115], [116] proposed to aggregate individual aesthetics into a single value and tested it in a user study. The measure included the aesthetic facets of edge crossing number, edge crossing angle, edge alignment, and edge length uniformity. A sample of 35 participants performed graph interpretation tasks to see if their performance agreed with the quality measure. Human performance and the measure agreed with large effect sizes on time, accuracy as well as subjective effort. This finding was later replicated in another study with a sample of 43 participants and a larger collection of graphs [117].

Summary: There seems to be a tendency for certain graph drawing aesthetics criteria. For example, the minimization of the number of edge bends and edge crossings [87], [89], [99], [112]–[114]. Also, the maximization of symmetries, maximization of angles between links [106]–[108], [111], and the fixing of nodes and links to a grid [89] are important criteria. For choosing algorithms focusing on aesthetics, it might be challenging since several of those stand in a trade-off behavior [90], [91], but using them in combination could also have benefits [109], [110], [115]–[117]. Also, the semantics of the graph dataset should be taken into account before choosing aesthetics criteria [93]–[96], but still fewer bends and link crossings are preferred [97], in particular, for following longer paths in a graph [98], i.e., making them as straight as possible. The geodesic-path tendency comes into play here [103]–[105].

c: CLUTTER REDUCTION TECHNIQUES

When the node-link visual metaphor is used to visually encode graph data, we soon reach a situation in which the viewer is represented a graph diagram containing the negative effect denoted by visual clutter [19].

Typically, a sophisticated layout algorithm is applied to produce node-link diagrams that are aesthetically pleasing, i.e., conform to a given list of aesthetic graph drawing criteria [86]. However, if the graph data gets too dense or several attributes are attached to the graph edges (as in multivariate graphs), further clutter reduction techniques come into play. Typically, edges are bundled in order to better reveal the structure of the graph and not the details about single edges.

In this line of research, Telea *et al.* [118] compared node-link diagrams with and without hierarchical edge bundling (see Figure 7). Five users were asked to give qualitative feedback. All users preferred the edge bundling technique for compound graphs that are larger than a few hundred nodes. They mention that edge bundled graphs are less cluttered, but that in node-link diagrams, paths can be traced more easily.

McGee and Dingliana [119] evaluated the user performance of 21 participants when using edge bundling by also varying graph density and graph size. Compound graphs were

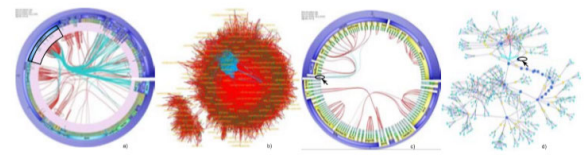


FIGURE 7. A study of node-link diagrams with and without edge bundling. Image reprinted from Telea *et al.* [118].

shown as hierarchical edge bundles by varying the bundling factor. Edge bundling was found to have a negative influence on the performance of path reading tasks for completion times and error rates. Higher-level cluster connectivity tasks, instead, could be solved significantly faster with edge bundling, but there was no significant effect on error rates.

A study by Bach *et al.* [120] compared four types of clutter reduction techniques in a user study with 15 participants. For path finding tasks, metro-style bundling, spatial edge bundling, power graphs, and confluent drawings were compared in terms of speed, accuracy, and user preferences. Quantitatively, power graphs performed best in terms of accuracy and speed, although metro-style bundling was also equivalently accurate. Spatial edge bundling was slowest. User preferences, however, indicated that power graphs were perceived as difficult to learn and there was more variability in perceptions of confidence in using them. In general, metro-style bundles and confluence graphs were perceived as most learnable and produced the highest level of confidence in graph interpretation.

A study by Dang *et al.* [121] compared six tree layouts that utilized hierarchical edge bundling. Classic, inverted radial, treemap, balloon, icicle, and cactus layouts were investigated on subtree identification and path tracing tasks. The 14 participants in the study were most accurate and fastest in identifying subtrees with the cactus and icicle layouts, treemap was the worst in terms of accuracy, and radial was worst in terms of speed. For the path tracing tasks, the cactus layout produced the fastest and most accurate performance, radial also produced good accuracy. User preference data also strongly favored the cactus layout for both tasks, and icicle layout was liked for the subtree identification task.

Graph clustering can also be regarded as a technique to reduce visual clutter since related vertices are spatially mapped next to each other, reducing link lengths and, consequently, the probability of link crossings. Archambault *et al.* [122] studied the readability of path-preserving clusterings versus no clustering in node-link graphs. If the graph is highly connected, clustering improves performance.

Summary: There seems to be a tendency to prefer edge bundling for large graphs [118], [121], but if path-related tasks have to be answered [118], [119], edge bundling is not the right choice. It seems to be better for cluster connectivity tasks [119], whereas also simple graph clustering can have benefits [122]. Metro-style bundling and power graphs are useful [120], while spatial edge bundling generates the

slowest response times [120]. From the hierarchical edge bundling approaches, cactus and icicle layouts seem to be the best ones [121], but the treemap layout is problematic.

3) SPECIAL PROPERTIES

There are also some graph visualization user studies that do not directly compare visual representations of graph dimensions or check different layout aspects for user performance, but they rather explore properties of the graph data. Section V-A3.a takes a look at special graph classes and the topology while Section V-A3.b explores edge properties.

a: GRAPH CLASSES AND TOPOLOGY

Burch *et al.* [123], [124] conducted an eye tracking study to find the best performer out of traditional, orthogonal, or radial tree diagrams. Moreover, they investigated visual task solution strategies of the study participants when finding the least common ancestor of a number of highlighted leaf nodes [124]. Traditional node-link tree diagrams with the root on top performed best. The analysis of the recorded eye movement data revealed more chaotic task solution strategies in the radial layouts and a cross-checking behavior that accounted for the almost doubled task completion time compared to traditional and orthogonal layouts.

Indented list and graph representations for ontologies were evaluated in an eye tracking study with 36 participants by Fu *et al.* [125]. The researchers came to the conclusion that indented lists perform better for information search tasks, whereas graphs are better for information processing tasks.

Symmetry in graphs is an important aesthetic criterion, and measures for determining it in graphs were compared to human judgments in a study by Welch and Kobourov [126]. Three different measures of symmetry were compared with a sample of 30 participants. Evidence was found to support the measure of reflective symmetry as being most in agreement with human judgment. Qualitative feedback also indicated that reflective symmetry was noticed more than other types of symmetry.

Symmetry perception in graphs was also investigated by de Luca *et al.* [127]. An online study using data from 56 participants found that horizontal symmetry was considered most important, followed by vertical and then translational symmetry. Adding rotation to all types of symmetry led to lower perceptions of symmetry. For rotational symmetry, it was generally found that a greater number of rotation axes (i.e., reflections) leads to a greater perception of symmetry. The exception to this was 4 axes, which may be perceived as more symmetric since it includes both horizontal and vertical symmetry.

The aesthetic properties of graph outlines were also studied by Carbon *et al.* [128] (see Figure 8). In an online study with a sample of 233 participants, it was found that curvature was important to perceptions of beauty as was lower levels of outline complexity. Higher levels of outline complexity, however, were associated with more interest. The results for judgments of solid shapes and graphs were in conformance,

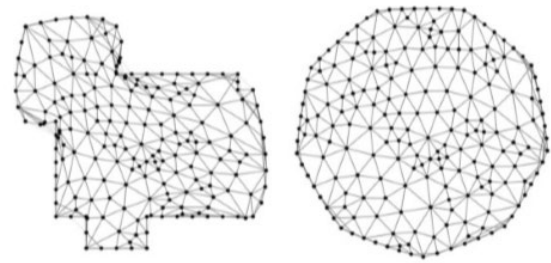


FIGURE 8. Evaluation of the power of shapes and properties of graph outlines. Unmodified image reprinted from Carbon *et al.* [128], © 2018 the authors, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>).

supporting the use of shape aesthetics research in guiding graph outlines.

User perception of graph meta-properties of different layouts was investigated by Soni *et al.* [129]. In two experiments, the perception of graph density and average local clustering in force-directed, circular, and multi-dimensional scaling layouts was determined. No differences were found between the layouts for the perception of graph density. The multi-dimensional scaling layout was found to produce a better perception of local clustering compared to the force-directed layout (the circular layout was not included).

The agreement between algorithmic detection of communities in social networks and user perceptions of such communities within their own social network was investigated by Lee and Archambault [130]. Twenty participants annotated their own social networks from a social media site. Six community finding algorithms, selected from the literature, were then let run on the networks. The user-annotated communities with automatically detected ones were compared and out of the three, the study identified the algorithm that was in closest agreement with the user annotations and was statistically better than the other algorithms.

The characteristics of similarity perception were also investigated by Ballweg *et al.* [131]. Using a card-sorting task with directed acyclic graphs, both quantitative and qualitative data were collected with a sample of 20 participants. Results indicated that depth, number of nodes per layer, and overall shape were important to similarity perception, but edge crossings were not. In determining similarity, it was found that participants used one of three strategies with up to 27 distinct graph factors that were mainly visual but also graph theoretical in nature.

Summary: There seems to be a tendency to hierarchical drawings in a top-to-bottom style with straight links [123], [124], also identifying a geodesic-path behavior by using eye tracking, for example, to explore search and information processing tasks [125]. Symmetry is an important feature [126], [127] as well as outlines [128] and similarity [131]. For meta-properties, it seems that a multi-dimensional scaling layout was best [129]. A comparison between algorithmic approaches and user annotations showed the benefits and drawbacks of automatic community detection [130].

b: EDGE PROPERTIES

Wattenberg [132] developed PivotGraph, a technique for representing multivariate graphs (each node has several attributes). The visualization technique was evaluated by showing it to 5 analysts who gave qualitative feedback. The biggest problem for the analysts was the fact that the topological graph properties were not easily observable.

Guo *et al.* [133] investigated the interaction between visual representations used for encoding edge attributes. Graphs that encoded both strength and certainty of edges were used. Strength was encoded using either width, hue, or saturation, and certainty was encoded using lightness, fuzziness, grain, and transparency. The discriminability of encodings was also varied by changing the amount of gradation between attribute levels (e.g., smaller or larger changes in width for increasing value). With a sample of 20 participants, a complex set of interactions was found between these variables on interpretation tasks involving search or heuristic judgment. It was found that some edge encoding styles interfered with each other, most notably lightness and hue interfered, as did width and fuzziness to reduce accuracy. It was also generally found that heuristic comparisons of overall strength or certainty between graphs were more effective than fine-grained judgments of searching for an edge with a specific value.

The encoding of uncertainty into edges using different visual representations was also investigated by Schwank *et al.* [134]. Edges used either dashes, stripes, blurring, or oscillating curves (waves) to represent different degrees of uncertainty. With a sample of 86 participants, some suggestive evidence was found that stripes performed worst for identifying certain edges, and waves performed worst for identifying uncertain edges in terms of speed and accuracy. Stripes were comparably slow but more accurate for uncertain edge identification. However, these differences do not seem to be reliable, and no statistical test was provided. Dashes overall received the best user preference ratings, but again no statistical test was provided.

Research by Bae *et al.* [135] investigating the interpretation of strength and certainty by edge properties showed difficulties for users. With a sample of 36 participants using tasks of indirect causality judgment, it was found that judgments of strength and certainty across two edges were inaccurate. Participants overestimated values and did so in a manner that was neither additive nor multiplicative.

An approach to displaying multivariate data on graph edges was studied by Schöffel *et al.* [136]. Bar charts of various types were displayed on the edges between nodes. A user evaluation with 89 participants found no reliable differences for efficiency or accuracy of graph interpretation tasks for the different types of bar charts investigated. User preference ratings, however, suggested that bars with bases on the edges were preferred to bars centered around the edges, equally sized bars were preferred to bars sized by edge length, and

bars oriented orthogonally to edges were preferred to a parallel orientation (however, no statistical tests were presented to verify these trends).

Augmented reality is a novel medium for displaying node-link graphs, and was subject to some preliminary research by Büschel *et al.* [137] (see Figure 9). The primary focus of the research was user acceptance testing for different link attribute encoding techniques on this medium. With a sample of 8 participants, it was found that color encoding was rated most aesthetically pleasing whilst blinking was least pleasing. Color was also rated highest for nominal data encoding, followed by static geometric techniques, and animated techniques were generally rated low. In contrast, encoding ordinal data had color rated second lowest, with animated techniques being preferred. A follow-up study [138] with 18 participants showed that even other edge variants based on shapes or geometry have some benefits in AR.



FIGURE 9. A node-link diagram from a study, showing links in different styles in an AR setting. Image reprinted from Büschel *et al.* [137].

Summary: There seems to be a tendency to simpler representations if many edge properties are contained in a graph dataset, for example, as bar charts [136]. An influence of edge encoding styles [133] was detected. Also, topological properties might be difficult to identify for multivariate graphs [132]. For example, to indicate uncertainty on edges, there are various ways, but a waved or striped style should be avoided [134]; instead, a dashed style might be the better option. In general, adding properties visually to edges can be a problem [135]. However, in an augmented reality environment, color might be a good visual attribute [137], whereas blinking might not be pleasing. However, shapes and geometry are feasible [138].

4) ADDITIONAL USER SUPPORT

More complex studies investigate the problem which influence visually enhanced graph representations have on user performance (Section V-A4.a). Moreover, interaction might be included in a user study, allowing a participant to change views on the represented graph or to navigate in it (Section V-A4.b).

a: ENRICHED GRAPH VISUALIZATIONS

Using additional visual references to enhance a graph visualization, either for accelerating and improving performances of graph interpretation tasks or for making it more memorable, has been studied by some researchers.

McGrath and Blythe [72] studied the effects of motion and spatial layout on the perception of a graph. Motion as a means to enrich a graph visualization was identified as having a positive effect on the 133 viewers' perception of change.

Information presented in animated or static forms was also investigated by Yoon *et al.* [139] for shortest path tasks. When animation was used to depict sequence information, it led to greater consideration of graph elements by 55 participants but had no effect on accuracy. When animation focused solely on elements relevant to the problem, compared to animating all elements, accuracy was improved, and participants attended less to irrelevant graph elements.

Multi-relational graphs are displayed by either color coding the links or not as an additional visual cue, which was investigated in a user study by Huang *et al.* [141]. Thirty students were recruited and were asked to perform tasks of different complexity levels. It was found that task performance varied with visual complexity of visualizations and task complexity. The impact of color coding was investigated by Karim *et al.* [140] for node-link diagrams (see Figure 10) by recruiting 35 participants. Blue colormaps seem to perform much better as well as viridis, producing fewer errors.

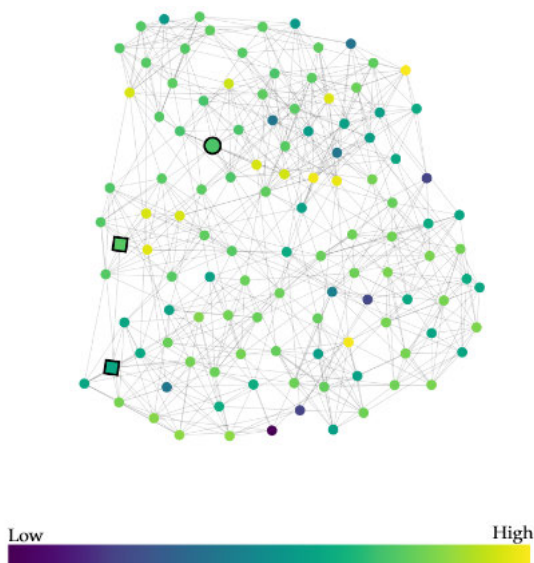


FIGURE 10. Evaluation of color in node-link diagrams. Unmodified image reprinted from Karim *et al.* [140], © 2019 the authors, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>).

The performance of difference maps in dynamic graphs was studied by Archambault *et al.* [48] by asking 25 participants. The difference map is a visual enrichment of a graph sequence to support change detection and the study found that difference maps outperformed other types of representations and were preferred by the participants.

Ghani and Elmqvist [142] found 16 volunteers for three experiments on exploring the performance of revisitation tasks in graphs by static spatial features. Size and color were the best node encoding techniques. For substrate encoding, landmarks and their combination performed well. Dunne and Shneiderman [143] introduced motif simplification, a technique to improve readability performance of networks by fan, connector, and clique glyphs. The effectiveness of the technique was demonstrated in case studies and a controlled experiment. Alper *et al.* [144] introduced the concept of additional grid or contour lines that they evaluated by asking 21 participants in total for three experiments about graphs with additional visual references such as grids or contour lines. Contour lines are beneficial for navigating large node-link diagrams.

Saket *et al.* [145] investigated if node, node-link, or node-link group diagrams performed best by asking 36 participants. Node-link group diagrams perform well for group-based diagrams. A subsequent study [146] investigated node groupings that were indicated either by node color or graph background color (i.e., map or Voronoi like). In an experiment with 40 participants, speed and accuracy advantages were found for grouping by background color for a range of graph interpretation tasks involving both nodes and node groups. Another subsequent study [147] with 17 participants using the same procedure but measuring enjoyment and engagement, found evidence that users prefer background color grouping over node color (although 2 out of 3 measures produced statistically equivocal results, all trended in that direction).

How group information might be displayed on top of node-link diagrams was also evaluated by Jianu *et al.* [148]. Node coloring, GMap, BubbleSets, and LineSets were compared by recruiting 788 subjects from the Amazon Mechanical Turk website. In this study, BubbleSets was shown to be the best technique when answering group membership tasks.

Visual enrichments by recording viewers' eye movements was evaluated by Okoe *et al.* [149]. Link dimming, highlighting, and increasing of saliency of nodes' neighborhoods were studied. The 12 participants were asked to answer tasks with and without gaze-enabled interaction. The results showed that gaze-enabled enriched graph visualization is beneficial for some situations.

Summary: There seems to be a tendency to use motion as a positive effect to indicate change in a graph [72] or animation for individual elements [139]. Also, difference maps as visual enrichment in dynamic graphs are beneficial to detect changes [48]. Typical visual variables like size and color [140], [142] are suitable visual enrichments, contour lines produce quite good effects [144], as well as motifs [143]. Moreover, group indications [145], [148], also shown by colored regions [146], are suitable features, but background color was found better than node color [147]. Even gaze-enabled visual enrichment was a promising concept [149]. However, visual complexities had a bad impact on performance [141].

b: GRAPH INTERACTIONS

There are also some papers that directly test if an interaction technique allowed in graph visualizations is useful and how well a user is performing when it is used in the study to solve certain tasks.

Ware and Bobrow developed and tested interactive motion highlighting techniques of graph elements over a series of experiments [13], [150], [151]. In the first study, neighborhood highlightings that used moving and static visual features were compared for node and edge search tasks with a sample of 13 participants [150]. It was found that motion improved the efficiency of visual search. This finding was extended by investigating graphs of greater size and complexity [13]. It was found that motion highlighting improved the effectiveness of visual search. When subgraphs needed to be searched for connections, contrasting static and motion highlighting improved both efficiency and effectiveness. A further study investigated motion and static highlighting on long-term memory performance of search task results [151]. Whilst long-term memory performance was poor, it was found that showing user interaction history information improved memory substantially by providing an effective cue to recall.

Skopik and Gutwin [152] also investigated interaction history highlighting on memory performance. Using an interactive fisheye distortion, 12 participants browsed a graph whilst memorizing the location of a specified node. It was found that focal point history highlighting improved memory accuracy and speed of recall compared to no highlighting for short-term memory.

Interactive features that provide temporal information that may reduce memory load were also investigated by Kondo *et al.* [63]. The authors tested clock-like element augmentations (see Figure 11) that displayed maps of element presence and absence over time as well allowing temporal manipulation using a temporal search task. In an experiment with 12 participants, these interactive features were compared to generic global time sliders for dynamic graphs. Quantitative analysis showed no differences between conditions. The authors argued that this was due to the eclectic use of controls that did not allow for a dissociation between time sliders and the special augments. Satisfaction ratings of the augments, however, were high.

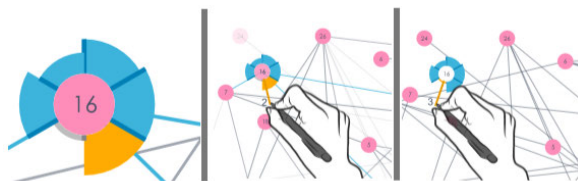


FIGURE 11. Evaluation of interactive features in the Glidgets approach. Image reprinted from Kondo *et al.* [63].

Another special interactive feature for graph search tasks was studied by Okoe *et al.* [149]. An eye tracking system was developed that intelligently detected user focus of attention

to highlight nodes, edges, and neighborhoods as well as filter out irrelevant graph elements. In an experiment with 12 participants, it was found that this highlighting feature improved the accuracy of search tasks for direct links between nodes. No reliable result was found for path searching tasks, which the authors argue was due to graph density reducing the enhancement effect of their highlighting feature.

Moscovich *et al.* [153] also investigated interactive features designed to facilitate path searching and graph navigation. Four navigation techniques were compared: pan and zoom, clickable bird's eye view (mini-map), link-sliding, and 'bring and go'. Pan and zoom and bird's eye view are standard navigation methods for maps and graphs. Link-sliding involved constrained viewport moving along a selected edge, and 'bring and go' moved all linked nodes to a selected node into the viewport. In an experiment with 12 participants, 'bring and go' was found to be the most efficient, effective, and satisfactory technique for graph navigation tasks.

For a graph generation and layout task, Purchase *et al.* [154]–[156] discuss differences in user interaction styles. In an initial usability study for the development of SketchNode, it was observed that users had different preferences for moving nodes throughout the creation process [154]. When users did free-drawings of graphs, they were less inclined to move nodes, place elements further apart, and organize things less orthogonally than when using a software diagramming approach with predefined elements [155]. Users, however, found free-drawing less cumbersome than navigating menus but preferred the neatness of predefined elements [156]. These observations may have implications for graph aesthetics.

The interface used for interacting with graphs may also affect interaction style and resultant graph aesthetics. Dwyer *et al.* [157] compared interaction patterns on a graph rearrangement task using mouse and touchscreen interfaces. A study involving 32 participants found that graphs rearranged with touch interactions were preferred more than those generated by mouse interactions. Touch interactions were characterized by a greater number of small manipulations but took no longer overall than mouse interactions.

Drogemuller *et al.* [158] evaluated two navigation techniques for graph visualizations in virtual reality (teleportation and one-handed flying) and other methods (two-handed flying and worlds in miniature). 25 participants' performance and effectiveness were measured for several tasks. Steering patterns (one-handed flying and two-handed flying) were faster and preferred by the participants for completing search tasks in comparison to teleportation.

Summary: There seems to be a tendency to show a user interaction history [151], [152], maybe also to support a visual search task in graphs by interactions (also in virtual reality [158]), in particular, motion highlighting [13], [150] or visual attention-based highlighting [149]. Special interactive augmentations [63] are suitable concepts, in particular, for graph navigation tasks, 'bring-and-go' was beneficial [153]. Node moving was less useful [154], rather pre-defined

locations were preferred [155], but free-drawing was less cumbersome than menu navigation [156]. Touch interaction was preferred over mouse interaction [157].

B. GRAPH MEMORABILITY

Apart from interpreting a graph visualization, we might also ask the question of how well a graph can be memorized. Memorization is important for tasks that need to keep certain aspects of a visualization in mind, because for a later re-inspection, the user might remember parts of this graph depiction to faster explore it when they are already familiar with it. Consequently, a body of research focuses on the specific challenge of investigating the user performance when trying to memorize a graph visualization. Section V-B1 describes memorability aspects in graph dynamics, whereas Section V-B2 focuses on the memorability of the graph structure.

1) MEMORABILITY IN GRAPH DYNAMICS

Sequence information takes a different form of memory than static graphs. Several studies were conducted on the memorability of dynamic graphs, mainly focusing on the effects of graph layout since node position calculations of layout algorithms can evolve over time as vertices and edges are added or removed.

The perception of animated node-link diagrams for dynamic graph visualization was researched by Ghani *et al.* [159]. Different metrics for dynamic graphs were tested for how they are perceived during a graph animation. A user study with 16 volunteers was conducted in which nodes and edges were added/deleted. After the animation, the participant had to reconstruct the order of these operations. A fixed layout was significantly more accurate than a corresponding force-directed layout. In a second study, 12 participants were recruited to also understand effects such as target separations and node speed. Lower correctness was achieved for high speed.

In a series of experiments, Archambault and Purchase [57], [160] investigated the memorability of dynamic graphs (see Figure 12). In the first experiment [160], 25 participants performed a recognition task on graphs that varied between constrained and free node movement between time slices. Node movement was animated and occurred due to nodes and edges being added or removed. These changes could thus affect their placement since the force-directed layout algorithm was

used. Results indicated that movement did not affect performance on a recognition task, however, participants preferred no movement. In a subsequent experiment [57] with 28 participants, it was found that node position recall was more efficient and effective when nodes were constrained, and participants' preference for this was also replicated. Another follow-up study [58] found that when nodes freely moved, animation was generally found to produce more accurate and faster recall of node position. However, these effects were only reliable for a smaller number of targets (1–3 for accuracy, 1 for speed).

Summary: There seems to be a tendency to non-animated dynamic graph visualizations [159]. Animation, in the form of node movement, was not preferred [160] and nodes should be constrained [57]. However, if animation is used, the speed should not be that high [159]. Animation should be used if only a small number of targets is present [58].

2) MEMORABILITY OF THE GRAPH STRUCTURE

The memorability of the graph structure was investigated by various authors. These studies generally manipulated various kinds of memory cues to find what best improves recognition and recall performance of graphs. Most studies focused on short-term memory, however, some involved long-term memory.

The effect of user interaction history information on memory performance was investigated by Skopic and Gutwin [152]. Using a system in which users could control the focal point of a fisheye distortion on a graph, it was found with 12 participants that node recollection speed was reduced. Interaction history in the form of highlighting of nodes that were within the focal point during memorization was compared to no highlighting with another group of 16 participants. User interaction history highlighting led to faster and more accurate recollection than no highlighting in a short-term memory task. A similar result was found for long-term memory after one week by Ware *et al.* [151]. In that study, no highlighting led to little or no recollection, but when interaction history was displayed, recollection improved substantially.

Another study involving distortions such as fisheye and graph memorability was conducted by Lam *et al.* [161]. This study found the degree of scaling, rotation, and fisheye distortions (rectangular and polar) that could be tolerated by participants before recognition performance suffered. It was found that graphs could be reduced to 20% of their original size, rotated up to 45 degrees, and a fisheye distortion factor up to 1–2 could be tolerated.

Perceptual organization may also influence graph memorability. A study by Marriott *et al.* [162] compared the memorability of graphs generated with layouts based on symmetry, node alignment, edge alignment, or no perceptual organizational principle (see Figure 13). Twenty-five participants memorized and then redrew graphs, and it was found that all graphs generated following perceptual organization



FIGURE 12. Evaluation of the mental map for memorizing the dynamics in a graph. Image reprinted from Archambault and Purchase [160], © 2012 IEEE.

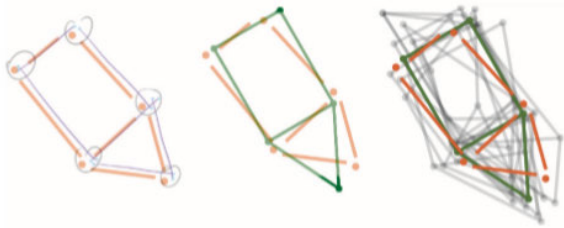


FIGURE 13. Comparison of memorability for graphs generated with different layouts. Image reprinted from Marriott *et al.* [162], © 2012 IEEE.

principles were recalled more easily. Graphs organized for symmetry were recalled best.

Visual spatial cues may also improve memorability according to an experiment by Ghani and Elmqvist [142]. In two node revisitation experiments, graph backgrounds and then node characteristics were varied to provide spatial cues. Sixteen participants attempted revisitation of graphs organized as a grid or a Voronoi diagram with either a matte color or photograph background, as well graphs with nodes that varied in color, size, or both depending on their coordinate position within 2D Cartesian space. Grid organization produced advantages for speed when the background was color, and accuracy when the background was a photograph. Nodes that varied in color and size produced an advantage for speed and accuracy compared to nodes that varied on a single attribute only. A second experiment with another sample of 16 participants compared grid, node, and landmarks (icons placed within grid locations) in all combinations. Those improved accuracy compared to no spatial information, and there was a trend toward grids improving the efficiency of revisitation.

Saket *et al.* [146] investigated both short-term and long-term recall of graph structure for graphs with grouped nodes. Groups were indicated either by node color or graph background color. With a sample of 40 participants, it was found that both immediate and long-term recall accuracy for graph structure were more accurate for the node groups that used background color.

A large crowdsourcing study by Okoe *et al.* [39] also investigated short-term recall and revisitation of node-link graphs and matrix representations. With a sample of 835 participants, they found that node-link diagrams outperformed matrices in terms of short-term recall. No differences were found for revisitation.

Summary: There seems to be a tendency to use interaction history for relocation tasks [152], also for long-term memory [151]. Also, changes in a graph could be tolerated to a certain degree [161]. If perceptual principles (like symmetry) were followed in a graph visualization, the graphs could be recalled more easily [162], similar findings were made for added spatial information [142] or background color [146]. Node-link diagrams outperformed matrices for a short-term recall [39].

C. GRAPH EXPRESSION AND CREATION

By graph expression and creation, the ability of a human user is understood to visually build a graph drawing either by hand or by a graph visualization tool. There are many studies that fall into this category that are usually not so concerned with the ability to create the drawing, but rather the manner in which it has been drawn. In this category, we found work on graph design from raw data (Section V-C1), and on how the graph is created by explicitly working with graph drawing and visualization tools (Section V-C2).

1) GRAPH VISUALIZATION BASED ON USER INTERVENTION

User-generated graphs can be affected by the format of the source information used to create them. Some studies directly compared how the source information format affects user-generated graphs. These studies often focus on the characteristics of the layouts generated by users.

A comparison between graphs produced from textual and diagrammatic source information was conducted by Tversky *et al.* [163]. A sample of 36 postgraduate students were given either textual descriptions or sequence diagrams (i.e., UML) to produce node-link diagrams of information systems. Results indicated that diagrammatic source information led to fewer deviations from the prototypical solution than the textual source. Evidence also suggested that the order of information presented in both sources affected the layout of elements in the graphs produced.

Purchase [164] also investigated the creation of node-link diagrams from two different formats of source information. As a follow-up of earlier work [155], [156], graphs created from a node-pair list were compared to graphs generated from a node adjacency table with a sample of 26 participants. Whilst graphs produced from node-pair lists showed a tendency toward a grid layout with horizontal and vertical edges, graphs produced from the adjacency list showed no such tendency.

Graph organization was also the focus of a study by Yu *et al.* [165]. This study asked participants to draw their own personal social networks with instructions to include various types of agents within them (see Figure 14). Analysis of the graphs produced by 74 participants indicated that graphs often used spatial organization to confer relationship closeness; familial relationships were nearer the participant's node than non-familial. A tendency was also observed for ancestors to be placed above the participant's node.

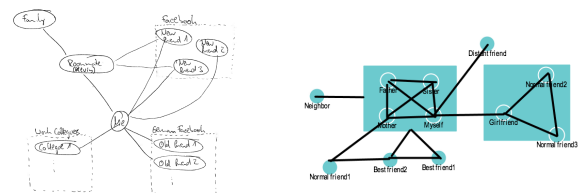


FIGURE 14. Evaluation of graph organization. Image provided by Yu *et al.*, from an earlier version of their paper [165].

Summary: There seems to be a tendency to use diagrammatic source information [163], which is better than a textual source. Graphs from a node-pair list typically led to a grid layout, whereas adjacency lists did not [155], [156], [164]. Spatial proximity is used to indicate relational closeness [165].

2) WORKING WITH GRAPH DRAWING TOOLS

Some studies investigated the use of node-link graph creation tools. These studies generally investigate user behaviors and preferences for the graph layouts they produce, as well as comparing various techniques and features.

User-generated layouts can include aesthetically pleasing features that are not found in algorithm-generated layouts and have prompted work into crowdsourcing as a layout method [166]. For example, a recent study by Singh *et al.* [167] looked into the possibility of using crowdsourcing to layout biological networks. Following a set of guidelines for biological network visualizations, expert-generated, crowd-generated, and algorithm-generated layouts were judged by an expert. In terms of the guidelines, there was no statistically significant difference between crowd and expert, although expert layouts were nominally better. Algorithm-generated layouts were the worst. Expert layouts were better in overall quality, but crowd layouts were generated much faster. In reviewing layouts, crowds were found to be fairly consistent in their appraisals, and reasonably consistent with expert appraisals.

Using a web-based graph editor, Van Ham and Rogowitz [168] compared user-generated layouts to force-directed layouts of the same graphs. Using 73 user-generated layouts, it was found that users avoided edge crossings more, had greater edge length variation, and organized nodes more symmetrically into horizontal and vertical orientations than the force-directed algorithm. Users also showed a strong preference for separating node clusters, often by containing them within a hull of edges.

A series of studies by Purchase *et al.* [154]–[156], [164] report on the development of SketchNode. This is a free-hand and diagramming tool for producing node-link graphs. A usability evaluation with the prototype [156] found that users preferred the free-hand mode for interaction but preferred the neater appearance of the diagramming mode's result. Participants also found an automatic layout feature for the diagramming mode useful. An initial study using the tool with a sample of 17 participants [154] found a consistent preference for avoiding link crossings as well as preferences for straight, horizontal, and vertical lines. In a follow-up study [155], it was found that diagramming mode led to graphs with tighter clusters of nodes with straighter edges than free-hand mode. Participants were more reluctant to move nodes in free-hand mode (even though this was possible) and used edges of more uniform length. These findings were later replicated and extended [164] by comparing preferences for user-generated layouts to force-directed and grid layouts. Immediately after completion, participants

preferred their own layouts, but changed their preferences to grid layouts after two weeks.

Following a similar evaluation approach by Purchase *et al.* [154], Lin and his colleagues conducted a series of user studies using an interactive drawing tool to investigate drawing behaviors of users and their preference and usage of aesthetics. In a study in which thirty participants were asked to draw clustered graphs from adjacency lists along with information about their clusters [169], user-sketched drawings (see Figure 15) and drawing strategies were examined. It was found that the cluster information helped participants to better organize their graphs. Graphs had fewer edge crossings and a substantial proportion were orthogonal and involved symmetry. Participants were also found to use more bends, especially within clusters.

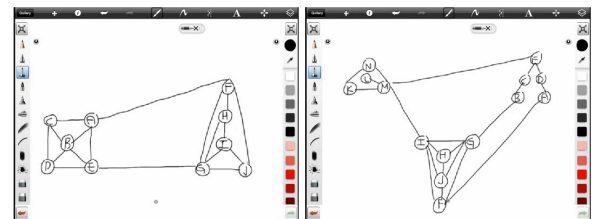


FIGURE 15. The aesthetics choices of users drawing clustered graphs. Image reprinted from Lin *et al.* [169], © 2016, with permission from Elsevier.

Under a different context, aesthetic choices when drawing graphs with additional symmetry information provided were later investigated by Lin *et al.* [170]. A sample of 30 participants drew graphs from adjacency lists on a tablet. In agreement with previous studies, it was found that participants tended to avoid edge crossings. This effect was particularly pronounced when participants were aware that the graph could be represented symmetrically. Node and edge overlapping, as well as curved edges were also generally avoided. When participants were told about symmetry, there was an increase in the degree of symmetry found in their drawings. This effect also persisted into a condition where no mention of symmetry was made. It also appeared that symmetry was given priority over edge crossing avoidance once participants were aware of it. A further finding of this study was that participants have different drawing strategies that persist despite changing goals (i.e., to create symmetrical graphs). Recently, Lin *et al.* [171] asked users in a study to draw vertex-weighted graphs and found that edge crossings were still the most important aesthetic and that grid-like drawings were preferred with vertex-weighted graphs.

Summary: There seems to be a tendency that the crowd can produce similar layouts as experts [167]. Users typically generate layouts without edge crossings [168], [170], avoid node and link overlap [170], do not use curved edges [170], focus on symmetries [168], and even use straight, vertical, and horizontal lines [154]. Free-hand mode is preferred [156] while people were more reluctant to move nodes in free-hand mode [155], [164]. Cluster information can help organize graphs [154], [169].

VI. OPEN RESEARCH CHALLENGES AND WHITE SPOTS

Although many empirical user studies exist in graph visualization, there is still a lot to be researched. Graph data comes in many data dimensions, i.e., vertices, edges, graph topological properties, or time. Moreover, due to the technical progress, more advanced display technologies and hardware devices are being constructed, which allow one to visually represent graph data from new perspectives, opening up new avenues for interaction.

A. STUDY DESIGNS

User studies in graph visualization have been conducted for more than 20 years now, and the aspects studied so far can literally be compared with the tip of the iceberg. There are many parameters that require various studies to fully understand the field. More techniques and layouts are developed every year, demanding additional user studies investigating further parameters to understand user performance with new designs.

Most of the studies recruited between 20 to 40 participants. If the parameter space is too large, many more people are needed, maybe in a crowdsourcing experiment that is typically conducted in an uncontrolled setup. However, such a large population study can help reduce the parameter space for a controlled user study that could then be conducted as a follow-up experiment. Whilst there have been a growing number of crowdsourcing experiments in recent years, these studies are still in a minority.

B. EVALUATION OF RECORDED DATA

The evaluation of traditional performance measures is often-times done in a detailed form. Also, the exploration of qualitative user feedback is presented in a summarized text giving extra insights into user preferences or difficulties participants had during the study.

Nowadays, we can observe an increase of eye tracking studies as a means of recording data of the visual attention that users pay when inspecting a presented graph visualization. The recording of such spatio-temporal data is not a problem anymore, but the analysis of the trajectory data is. Typically, analyses on this kind of data do not allow one to derive visual task solution strategies that the user applied when solving the task. There should be more research in this direction, from which not only the graph visualization community could benefit, but also other disciplines dealing with recorded eye movement data.

C. INDIVIDUAL DIFFERENCES

We visualize graphs for end-users to gain insights into the data and/or communicate the insights to others. However, there are individual differences between people. Different users have different capabilities and strategies in perceiving and processing information. And these strategies can vary when the amount of information to be processed increases to some point. To evaluate how good a visualization is in

conveying data information to end-users, it is important to take into consideration the characteristics of the target audience. This includes: cognitive ability, cognitive style, age, gender, ethics background, and memory span, to name a few. In general, although occasionally some user studies considered one or two individual difference factors (e.g., [172]), a systematic evaluation of individual differences is still uncommon. Therefore, there is a need to research further in this space so that pros and cons of evaluated visualizations can be fully understood in relation to their target users.

D. NEW AESTHETICS

Current aesthetics, such as minimum crossing number and uniform edge lengths, have been well researched by researchers in the graph drawing community. They were originally proposed in terms of visual graph layout features by visualization and algorithm designers mainly for the purpose of making graph drawings visually pleasing. Although research has shown that graph drawings can be effective if they conform to the layout-based aesthetic criteria, the impact of different aesthetics on the effectiveness of graph drawings varies.

As the ultimate goal of graph visualization is to convey underlying graph structural information embedded in the drawing to end-users, it is reasonable to propose or derive aesthetics based on graph structural features and end-user graph drawing behaviors [173]. Future aesthetics that are to be used to judge the quality of visualizations should be generated and formed from the perspective of end-users, and attempts have been made in recent research in this direction [174]. For example, a graph was drawn so that the physical distance of any two nodes reflects the length of their theoretical shortest path; end-users were invited to participate in experiments in which they were asked to draw graphs as they wish so that their drawing behavior and preference could be observed and new aesthetics could be derived.

E. NEW MEASUREMENTS AND EVALUATION METHODOLOGIES

Individual aesthetics are often used to evaluate the extents to which a drawing meets individual drawing rules. However, when it comes to human graph comprehension, it is the overall quality that is more relevant. Overall quality can be considered as a result of interactions between individual aesthetics. However, how to accurately measure the overall quality of a visualization is still unknown.

Further, most user studies are conducted in laboratory settings with university students. Although controlled laboratory studies give us useful insights into the quality of visualization, the findings can be limited in revealing the true value of a visualization used in an environment that is very different from an academic laboratory and by users who are not properly represented by students. Therefore, new evaluation methodologies are needed to reflect the true purposes of evaluated visualizations.

VII. CONCLUSION

In this article, we presented the state of the art in empirical user evaluation of graph visualizations. Graph data visualization is a challenging discipline due to the many data dimensions, features, and display opportunities that are more or less the focus of user evaluations. An overview of the already found insights in user perception and performance when answering tasks was missing before, making it hard to evaluate further aspects of graph visualization. Therefore, we summarized existing work and tried to identify white spots in research.

While there is a large body of work on graph interpretation, in particular, on graph layouts and the aesthetics of those drawings, as well as dynamic graph visualization evaluation, only a few approaches exist on the memorization of graph visualizations and also on how people create graphs, i.e., how graphs are taught and how well people perform when learning graph visualizations.

By systematically filling our research database and by categorizing the papers, we were able to structure the field of graph visualization evaluation. This process demanded to search for papers in nearly any visualization discipline since graphs are a wide-spread data structure covering many areas. By studying the evolution in this field, we can say that research in graph visualization evaluation is far away from coming to an end. Novel evaluation techniques such as eye tracking provide additional technologies to deeper dive into the details of human cognition and perception. However, it also poses additional challenges for the evaluation of the recorded data, which has a spatio-temporal nature. Also, the introduction of new graph visualizations, systems, and interaction techniques is likely to produce new user evaluations that will have to be integrated into a future state-of-the-art report.

REFERENCES

- [1] F. Beck, M. Burch, S. Diehl, and D. Weiskopf, "A taxonomy and survey of dynamic graph visualization," *Comput. Graph. Forum*, vol. 36, no. 1, pp. 133–159, Jan. 2017.
- [2] T. von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. van Wijk, J.-D. Fekete, and D. W. Fellner, "Visual analysis of large graphs: State-of-the-art and future research challenges," *Comput. Graph. Forum*, vol. 30, no. 6, pp. 1719–1749, Sep. 2011.
- [3] V. Yeghouchian, D. Archambault, S. Diehl, T. Dwyer, K. Klein, H. C. Purchase, and H.-Y. Wu, "Exploring the limits of complexity: A survey of empirical studies on graph visualisation," 2018, *arXiv:1809.00270*. [Online]. Available: <http://arxiv.org/abs/1809.00270>
- [4] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry, "Task taxonomy for graph visualization," in *Proc. AVI Workshop BEyond Time Errors Novel Eval. Methods Inf. Vis.*, 2006, pp. 1–5.
- [5] L. Euler, "Solutio problematis ad geometriam situs pertinentis," *Commentarii Academiae Scientiarum Petropolitanae*, vol. 8, pp. 128–140, 1741.
- [6] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale, "Empirical studies in information visualization: Seven scenarios," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 9, pp. 1520–1536, Sep. 2012.
- [7] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller, "A systematic review on the practice of evaluating visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2818–2827, Dec. 2013.
- [8] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale, "Benchmarks for graph embedding evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 9, pp. 1520–1536, 2011.
- [9] S. Carpendale, "Information visualization," in *Evaluating Information Visualizations*, A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, Eds. Berlin, Germany: Springer-Verlag, 2008, pp. 19–45.
- [10] G. D. Battista, P. Eades, R. Tamassia, and I. G. Tollis, *Graph Drawing: Algorithms for the Visualization of Graphs*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [11] M. Kaufmann and D. Wagner, *Drawing Graphs, Methods Models* (Lecture Notes in Computer Science), vol. 2025. Berlin, Germany: Springer-Verlag, 2001.
- [12] P. Erdős and A. Rényi, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.
- [13] C. Ware and R. Bobrow, "Supporting visual queries on medium-sized node-link diagrams," *Inf. Vis.*, vol. 4, no. 1, pp. 49–58, 2005.
- [14] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
- [15] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [16] D. Holten, P. Isenberg, J. J. van Wijk, and J.-D. Fekete, "An extended evaluation of the readability of tapered, animated, and textured directed-edge representations in node-link graphs," in *Proc. IEEE Pacific Vis. Symp.*, Mar. 2011, pp. 195–202.
- [17] K. Sugiyama, S. Tagawa, and M. Toda, "Methods for visual understanding of hierarchical system structures," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-11, no. 2, pp. 109–125, Feb. 1981.
- [18] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Softw. Pract. Exper.*, vol. 21, no. 11, pp. 1129–1164, Nov. 1991.
- [19] R. Rosenholtz, Y. Li, J. Mansfield, and Z. Jin, "Feature congestion: A measure of display clutter," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2005, pp. 761–770.
- [20] M. Bar and M. Neta, "Humans prefer curved visual objects," *Psychol. Sci.*, vol. 17, no. 8, pp. 645–648, 2006.
- [21] M. Pohl, M. Schmitt, and S. Diehl, "Comparing the readability of graph layouts using eyetracking and task-oriented analysis," in *Proc. Euro-graph. Workshop Comput. Aesthetics*, 2009, pp. 49–56.
- [22] U. Föbmeier and M. Kaufmann, "Drawing high degree graphs with low bend numbers," in *Proc. Symp. Graph Drawing*, 1995, pp. 254–266.
- [23] D. Holten, P. Isenberg, J.-D. Fekete, and J. van Wijk, "Performance evaluation of tapered, curved, and animated directed-edge representations in node-link graphs," Res. Rep. HAL-00696823, 2010. [Online]. Available: <https://hal.inria.fr/hal-00696823>
- [24] D. Holten and J. J. van Wijk, "A user study on visualizing directed edges in graphs," in *Proc. Int. Conf. Hum. Factors Comput. Syst.*, 2009, pp. 2299–2308.
- [25] M. Okoe and R. Jianu, "GraphUnit: Evaluating interactive graph visualizations using crowdsourcing," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 451–460, Jun. 2015.
- [26] K. Xu, C. Rooney, P. Passmore, D. Ham, and P. H. Nguyen, "A user study on curved edges in graph visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2449–2456, Oct. 2012.
- [27] M. Couch, "Effects of curved lines on force-directed graphs," M.S. thesis, Dept. Comput. Inf. Sci., East Tennessee State Univ., Johnson City, TN, USA, 2013. [Online]. Available: <https://dc.etsu.edu/honors/64>
- [28] W. Huang, P. Eades, S.-H. Hong, and H. B.-L. Duh, "Effects of curves on graph perception," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Apr. 2016, pp. 199–203.
- [29] A. Rusu, A. J. Fabian, R. Jianu, and A. Rusu, "Using the gestalt principle of closure to alleviate the edge crossing problem in graph drawings," in *Proc. 15th Int. Conf. Inf. Vis.*, Jul. 2011, pp. 488–493.
- [30] M. Burch, C. Vehlou, N. Konevtsova, and D. Weiskopf, "Evaluating partially drawn links for directed graph edges," in *Proc. Graph Drawing*, 2011, pp. 226–237.
- [31] M. Burch, "A user study on judging the target node in partial link drawings," in *Proc. 21st Int. Conf. Inf. Visualisation (IV)*, Jul. 2017, pp. 199–204.
- [32] M. Sathyanarayanan and D. Pirozzi, "Social network visualization: Does partial edges affect user comprehension?" in *Proc. 9th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2017, pp. 570–575.
- [33] N. Henry and J. Fekete, "MatLink: Enhanced matrix visualization for analyzing social networks," in *Proc. Conf. Human-Comput. Interact. (INTERACT)*, 2007, pp. 288–302.
- [34] N. Henry, J.-D. Fekete, and M. J. McGuffin, "NodeTrix: A hybrid visualization of social networks," *IEEE Trans. Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1302–1309, Nov. 2007.

- [35] M. Ghoniem, J. Fekete, and P. Castagliola, "A comparison of the readability of graphs using node-link and matrix-based representations," in *Proc. IEEE Symp. Inf. Vis.*, Oct. 2004, pp. 17–24.
- [36] J. Bertin, *Semiology of Graphics: Diagrams, Networks, Maps*. Redlands, CA, USA: Esri Press, 2010.
- [37] M. Hlawatsch, M. Burch, and D. Weiskopf, "Visual adjacency lists for dynamic graphs," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 11, pp. 1590–1603, Nov. 2014.
- [38] M. Ghoniem, J.-D. Fekete, and P. Castagliola, "On the readability of graphs using node-link and matrix-based representations: A controlled experiment and statistical analysis," *Inf. Vis.*, vol. 4, no. 2, pp. 114–135, Jun. 2005.
- [39] M. Okoe, R. Jianu, and S. Kobourov, "Node-link or adjacency matrices: Old question, new insights," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 10, pp. 2940–2952, Oct. 2019.
- [40] R. Keller, C. M. Eckert, and P. J. Clarkson, "Matrices or node-link diagrams: Which visual representation is better for visualising connectivity models?" *Inf. Vis.*, vol. 5, no. 1, pp. 62–76, Mar. 2006.
- [41] B. Alper, B. Bach, N. H. Riche, T. Isenberg, and J. Fekete, "Weighted graph comparison techniques for brain connectivity analysis," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst.*, 2013, pp. 483–492.
- [42] M. McBride and M. Caldara, "The efficacy of tables versus graphs in disrupting dark networks: An experimental study," *Social Netw.*, vol. 35, no. 3, pp. 406–422, Jul. 2013.
- [43] F. Beck, M. Burch, and S. Diehl, "Towards an aesthetic dimensions framework for dynamic graph visualisations," in *Proc. Int. Conf. Inf. Vis.*, 2009, pp. 592–597.
- [44] F. Beck, M. Burch, and S. Diehl, "Matching application requirements with dynamic graph visualization profiles," in *Proc. 17th Int. Conf. Inf. Vis.*, Jul. 2013, pp. 11–18.
- [45] S. S. Bridgeman and R. Tamassia, "A user study in similarity measures for graph drawing," in *Proc. Int. Symp. Graph Drawing*, 2000, pp. 19–30.
- [46] S. Bridgeman and R. Tamassia, "Difference metrics for interactive orthogonal graph drawing algorithms," *J. Graph Algorithms Appl.*, vol. 4, no. 3, pp. 47–74, 2000.
- [47] S. Bridgeman and R. Tamassia, "A user study in similarity measures for graph drawing," *J. Graph Algorithms Appl.*, vol. 6, no. 3, pp. 225–254, 2002.
- [48] D. Archambault, H. C. Purchase, and B. Pinaud, "Difference map readability for dynamic graphs," in *Proc. Int. Symp. Graph Drawing*, 2010, pp. 50–61.
- [49] L. Zaman, A. Kalra, and W. Stürzlinger, "The effect of animation, dual view, difference layers, and relative re-layout in hierarchical diagram differencing," in *Proc. Graph. Interface Conf.*, 2011, pp. 183–190.
- [50] P. Saraiya, P. Lee, and C. North, "Visualization of graphs with associated timeseries data," in *Proc. IEEE Symp. Inf. Vis.*, Oct. 2005, pp. 225–232.
- [51] H. C. Purchase, E. Hoggan, and C. Görg, "How important is the mental map? An empirical investigation of a dynamic graph layout algorithm," in *Proc. Int. Symp. Graph Drawing*, 2007, pp. 262–273.
- [52] H. C. Purchase and A. Samra, "Extremes are better: Investigating mental map preservation in dynamic graphs," in *Proc. Diagrammatic Represent. Inference (Diagrams)*, 2008, pp. 60–73.
- [53] P. Saffrey and H. Purchase, "The mental map versus static aesthetic compromise in dynamic graphs: A user study," in *Proc. 9th Australas. User Interface Conf. (AUIC)*, 2008, pp. 85–93.
- [54] D. Archambault, H. Purchase, and B. Pinaud, "Animation, small multiples, and the effect of mental map preservation in dynamic graphs," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 4, pp. 539–552, Apr. 2011.
- [55] D. Archambault and H. C. Purchase, "The 'map' in the mental map: Experimental results in dynamic graph drawing," *Int. J. Hum.-Comput. Stud.*, vol. 71, no. 11, pp. 1044–1055, 2013.
- [56] D. Archambault and H. C. Purchase, "On the application of experimental results in dynamic graph drawing," in *Proc. 4th Int. Workshop Euler Diagrams 1st Int. Workshop Graph Vis. Pract.*, 2014, pp. 73–77.
- [57] D. Archambault and H. C. Purchase, "Mental map preservation helps user orientation in dynamic graphs," in *Proc. Int. Symp. Graph Drawing*, 2012, pp. 475–486.
- [58] D. Archambault and H. C. Purchase, "Can animation support the visualisation of dynamic graphs?" *Inf. Sci.*, vol. 330, pp. 495–509, Feb. 2016.
- [59] D. Archambault and H. C. Purchase, "On the effective visualisation of dynamic attribute cascades," *Inf. Vis.*, vol. 15, no. 1, pp. 51–63, Jan. 2016.
- [60] I. Boyandin, E. Bertini, and D. Lalanne, "A qualitative study on the exploration of temporal changes in flow maps with animation and small-multiples," *Comput. Graph. Forum*, vol. 31, no. 3, pp. 1005–1014, Jun. 2012.
- [61] G. D. Rey and S. Diehl, "Controlling presentation speed, labels, and tooltips in interactive dynamic graphs," *J. Media Psychol. Theories, Methods, Appl.*, vol. 22, no. 4, pp. 160–170, 2010.
- [62] L. Shi, C. Wang, Z. Wen, H. Qu, C. Lin, and Q. Liao, "1.5D egocentric dynamic network visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 21, no. 5, pp. 624–637, May 2015.
- [63] B. Kondo, H. Mehta, and C. Collins, "Glidgets: Interactive glyphs for exploring dynamic graphs," in *Proc. Posters IEEE VIS*, 2014.
- [64] B. Bach, E. Pietriga, and J. Fekete, "GraphDiaries: Animated transitions and temporal navigation for dynamic networks," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 5, pp. 740–754, May 2014.
- [65] S. Rufiange and M. J. McGuffin, "DiffAni: Visualizing dynamic graphs with a hybrid of difference maps and animation," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2556–2565, Dec. 2013.
- [66] M. Burch, F. Bott, F. Beck, and S. Diehl, "Cartesian vs. radial—A comparative evaluation of two visualization tools," in *Proc. Int. Symp. Adv. Vis. Comput.*, 2008, pp. 151–160.
- [67] M. Burch, F. Beck, and S. Diehl, "Timeline trees: Visualizing sequences of transactions in information hierarchies," in *Proc. Work. Conf. Adv. Vis. Interfaces*, 2008, pp. 75–82.
- [68] M. Burch and S. Diehl, "TimeRadarTrees: Visualizing dynamic compound digraphs," *Comput. Graph. Forum*, vol. 27, no. 3, pp. 823–830, May 2008.
- [69] J. Blythe, C. McGrath, and D. Krackhardt, "The effect of graph layout on inference from social network data," in *Proc. Symp. Graph Drawing*, 1995, pp. 40–51.
- [70] C. McGrath, J. Blythe, and D. Krackhardt, "Seeing groups in graph layouts," *Connections*, vol. 19, no. 2, pp. 22–29, 1996.
- [71] C. McGrath, J. Blythe, and D. Krackhardt, "The effect of spatial arrangement on judgments and errors in interpreting graphs," *Social Netw.*, vol. 19, no. 3, pp. 223–242, Aug. 1997.
- [72] C. McGrath and J. Blythe, "Do you see what I want you to see? The effects of motion and spatial layout on viewers' perceptions of graph structure," *J. Social Struct.*, vol. 5, no. 2, 2004.
- [73] S. Chaturvedi, C. Dunne, Z. Ashktorab, R. Zachariah, and B. Shneiderman, "Group-in-a-box meta-layouts for topological clusters and attribute-based groups: Space-efficient visualizations of network communities and their ties," *Comput. Graph. Forum*, vol. 33, no. 8, pp. 52–68, Dec. 2014.
- [74] W. Meulemans and A. Schulz, "A tale of two communities: Assessing homophily in node-link diagrams," in *Graph Drawing and Network Visualization*. Springer, 2015, pp. 489–501.
- [75] E. Dengler and W. B. Cowan, "Human perception of laid-out graphs," in *Proc. Int. Symp. Graph Drawing*, 1998, pp. 441–443.
- [76] H. C. Purchase, "Performance of layout algorithms: Comprehension, not computation," *J. Vis. Lang. Comput.*, vol. 9, no. 6, pp. 647–657, Dec. 1998.
- [77] H. C. Purchase, L. Colpoys, M. McGill, and D. A. Carrington, "UML collaboration diagram syntax: An empirical study of comprehension," in *Proc. 1st Int. Workshop Visualizing Softw. Understand. Anal.*, 2002, pp. 13–22.
- [78] H. C. Purchase, R. Welland, M. McGill, and L. Colpoys, "Comprehension of diagram syntax: An empirical study of entity relationship notations," *Int. J. Hum.-Comput. Stud.*, vol. 61, no. 2, pp. 187–203, Aug. 2004.
- [79] C. Körner, "Sequential processing in comprehension of hierarchical graphs," *Appl. Cognit. Psychol.*, vol. 18, no. 4, pp. 467–480, 2004.
- [80] C. Körner, "Eye movements reveal distinct search and reasoning processes in comprehension of complex graphs," *Appl. Cognit. Psychol.*, vol. 25, no. 6, pp. 893–905, Nov. 2011.
- [81] W. Didimo, E. M. Kornaropoulos, F. Montecchiani, and I. G. Tollis, "A visualization framework and user studies for overloaded orthogonal drawings," *Comput. Graph. Forum*, vol. 37, no. 1, pp. 288–300, Feb. 2018.
- [82] W. Huang and P. Eades, "How people read graphs," in *Proc. Asia-Pacific Symp. Inf. Vis.*, 2005, pp. 51–58.
- [83] R. Bourqui, H. C. Purchase, and F. Jourdan, "Domain specific vs generic network visualization: An evaluation with metabolic networks," in *Proc. 12th Australas. User Interface Conf.*, 2011, pp. 9–18.

- [84] H. C. Purchase, J. Hamer, M. Nöllenburg, and S. G. Kobourov, "On the usability of Lombardi graph drawings," in *Proc. Int. Symp. Graph Drawing*, 2012, pp. 451–462.
- [85] J. Q. Dawson, T. Munzner, and J. McGrenere, "A search set model of path tracing in graphs," *Inf. Vis.*, vol. 14, no. 4, pp. 308–338, 2014.
- [86] C. Bennett, J. Ryall, L. Spalteholz, and A. Gooch, "The aesthetics of graph visualization," in *Proc. Eurographics Workshop Comput. Aesthetics Graph., Vis. Imag.*, 2007, pp. 57–64.
- [87] H. C. Purchase, R. F. Cohen, and M. I. James, "Validating graph drawing aesthetics," in *Proc. Symp. Graph Drawing*, 1995, pp. 435–446.
- [88] H. C. Purchase, R. F. Cohen, and M. I. James, "An experimental study of the basis for graph drawing algorithms," *ACM J. Exp. Algorithmics*, vol. 2, p. 4, 1997.
- [89] H. C. Purchase, "Which aesthetic has the greatest effect on human understanding?" in *Proc. Int. Symp. Graph Drawing*, 1997, pp. 248–261.
- [90] H. C. Purchase, D. A. Carrington, and J. Alder, "Experimenting with aesthetics-based graph layout," in *Proc. Int. Conf. Theory Appl. Diagrams*, 2000, pp. 498–501.
- [91] H. C. Purchase, "Effective information visualisation: A study of graph drawing aesthetics and algorithms," *Interacting Comput.*, vol. 13, no. 2, pp. 147–162, Dec. 2000.
- [92] H. Eichelberger, "Aesthetics of class diagrams," in *Proc. 1st Int. Workshop Visualizing Softw. Understand. Anal.*, 2002, pp. 23–31.
- [93] H. C. Purchase, J. Alder, and D. A. Carrington, "User preference of graph layout aesthetics: A UML study," in *Proc. Int. Symp. Graph Drawing*, 2000, pp. 5–18.
- [94] H. C. Purchase, J.-A. Alder, and D. Carrington, "Graph layout aesthetics in UML diagrams: User preferences," *J. Graph Algorithms Appl.*, vol. 6, no. 3, pp. 255–279, 2002.
- [95] H. C. Purchase, M. McGill, L. Colpoys, and D. A. Carrington, "Graph drawing aesthetics and the comprehension of UML class diagrams: An empirical study," in *Proc. Australas. Symp. Inf. Vis.*, 2001, pp. 129–137.
- [96] H. C. Purchase, D. A. Carrington, and J. Alder, "Empirical evaluation of aesthetics-based graph layout," *Empirical Softw. Eng.*, vol. 7, no. 3, pp. 233–255, 2002.
- [97] D. Carrington, L. Colpoys, M. McGill, and H. Purchase, "Experimenting with layout and notation in UML diagrams," in *Proc. Int. Symp. Empirical Softw. Eng.*, 2002.
- [98] C. Ware, H. Purchase, L. Colpoys, and M. McGill, "Cognitive measurements of graph aesthetics," *Inf. Vis.*, vol. 1, no. 2, pp. 103–110, Jun. 2002.
- [99] C. Körner and D. Albert, "Speed of comprehension of visualized ordered sets," *J. Experim. Psychology: Appl.*, vol. 8, no. 1, pp. 57–71, 2002.
- [100] W. Huang, S. Hong, and P. Eades, "Layout effects on sociogram perception," in *Proc. Int. Symp. Graph Drawing*, 2005, pp. 262–273.
- [101] W. Huang, S.-H. Hong, and P. Eades, "Effects of sociogram drawing conventions and edge crossings in social network visualization," *J. Graph Algorithms Appl.*, vol. 11, no. 2, pp. 397–429, 2007.
- [102] W. Huang, S.-H. Hong, and P. Eades, "How people read sociograms: A questionnaire study," in *Proc. Asia-Pacific Symp. Inf. Vis.*, 2006, pp. 199–206.
- [103] W. Huang, "Using eye tracking to investigate graph layout effects," in *Proc. 6th Int. Asia-Pacific Symp. Vis.*, Feb. 2007, pp. 97–100.
- [104] W. Huang, "An eye tracking study into the effects of graph layout," *CoRR*, vol. abs/0810.4431, Oct. 2008.
- [105] W. Huang, P. Eades, and S.-H. Hong, "A graph reading behavior: Geodesic-path tendency," in *Proc. IEEE Pacific Vis. Symp.*, Apr. 2009, pp. 137–144.
- [106] W. Huang, S.-H. Hong, and P. Eades, "Effects of crossing angles," in *Proc. IEEE Pacific Vis. Symp.*, Mar. 2008, pp. 41–46.
- [107] W. Huang, P. Eades, and S.-H. Hong, "Larger crossing angles make graphs easier to read," *J. Vis. Lang. Comput.*, vol. 25, no. 4, pp. 452–465, Aug. 2014.
- [108] W. Huang, P. Eades, and S.-H. Hong, "Beyond time and error: A cognitive approach to the evaluation of graph drawings," in *Proc. Beyond Time Errors, Novel Eval. Methods Inf. Vis. (BELIV)*, 2008, p. 3.
- [109] W. Huang, P. Eades, S.-H. Hong, and C.-C. Lin, "Improving force-directed graph drawings by making compromises between aesthetics," in *Proc. IEEE Symp. Vis. Lang. Hum.-Centric Comput.*, Sep. 2010, pp. 176–183.
- [110] W. Huang, P. Eades, S.-H. Hong, and C.-C. Lin, "Improving multiple aesthetics produces better graph drawings," *J. Vis. Lang. Comput.*, vol. 24, no. 4, pp. 262–272, Aug. 2013.
- [111] W. Huang, M. Huang, and C.-C. Lin, "Aesthetic of angular resolution for node-link diagrams: Validation and algorithm," in *Proc. IEEE Symp. Vis. Lang. Hum.-Centric Comput.*, Sep. 2011, pp. 213–216.
- [112] W. Huang and M. L. Huang, "Exploring the relative importance of crossing number and crossing angle," in *Proc. Int. Symp. Vis. Inf. Commun.*, 2010, p. 10.
- [113] W. Huang and M. Huang, "Exploring the relative importance of number of edge crossings and size of crossing angles: A quantitative perspective," *Int. J. Adv. Intell.*, vol. 3, no. 1, pp. 25–42, 2011.
- [114] W. Huang, "Establishing aesthetics based on human graph reading behavior: Two eye tracking studies," *Pers. Ubiquitous Comput.*, vol. 17, no. 1, pp. 93–105, Jan. 2013.
- [115] W. Huang, C. Lin, and M. L. Huang, "An aggregation-based approach to quality evaluation of graph drawings," in *Proc. Int. Symp. Vis. Inf. Commun. Interact. (VINCI)*, 2012, pp. 110–113.
- [116] W. Huang, M. L. Huang, and C.-C. Lin, "Evaluating overall quality of graph visualizations based on aesthetics aggregation," *Inf. Sci.*, vol. 330, pp. 444–454, Feb. 2016.
- [117] W. Huang, M. Zhu, M. L. Huang, and H. B.-L. Duh, "Evaluating overall quality of dynamic network visualizations," in *Cooperative Design, Visualization, and Engineering*. Cham, Switzerland: Springer, 2016, pp. 157–162.
- [118] A. Telea, O. Ersoy, H. Hoogendorp, and D. Reniers, "Comparison of node-link and hierarchical edge bundling layouts: A user study," in *Visualization and Monitoring of Network Traffic* (Dagstuhl Seminar Proceedings), D. A. Keim, A. Pras, J. Schönwälder, and P. C. Wong, Eds. Dagstuhl, Germany: Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2009.
- [119] F. McGee and J. Dingliana, "An empirical study on the impact of edge bundling on user comprehension of graphs," in *Proc. Int. Work. Conf. Adv. Vis. Interface*, 2012, pp. 620–627.
- [120] B. Bach, N. H. Riche, C. Hurter, K. Marriott, and T. Dwyer, "Towards unambiguous edge bundling: Investigating confluent drawings for network visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 541–550, Jan. 2017.
- [121] T. Dang, P. Murray, R. Etemadpour, and A. G. Forbes, "A user study of techniques for visualizing structure and connectivity in hierarchical datasets," in *Proc. VOILA ISWC*, 2017, pp. 45–59.
- [122] D. Archambault, H. C. Purchase, and B. Pinaud, "The readability of path-preserving clusterings of graphs," *Comput. Graph. Forum*, vol. 29, no. 3, pp. 1173–1182, Aug. 2010.
- [123] M. Burch, N. Konevtsova, J. Heinrich, M. Höferlin, and D. Weiskopf, "Evaluation of traditional, orthogonal, and radial tree diagrams by an eye tracking study," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2440–2448, Dec. 2011.
- [124] M. Burch, G. Andrienko, N. Andrienko, M. Höferlin, M. Raschke, and D. Weiskopf, "Visual task solution strategies in tree diagrams," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Feb. 2013, pp. 169–176.
- [125] B. Fu, N. F. Noy, and M.-A. Storey, "Eye tracking the user experience—An evaluation of ontology visualization techniques," *Semantic Web-Interoperability, Usability, Applicability*, vol. 8, no. 1, pp. 23–41, 2017.
- [126] E. Welch and S. Kobourov, "Measuring symmetry in drawings of graphs," *Comput. Graph. Forum*, vol. 36, no. 3, pp. 341–351, Jun. 2017.
- [127] F. De Luca, S. Kobourov, and H. Purchase, "Perception of symmetries in drawings of graphs," 2018, *arXiv:1808.01031*. [Online]. Available: <http://arxiv.org/abs/1808.01031>
- [128] C.-C. Carbon, T. Mchedlidze, M. H. Raab, and H. Wächter, "The power of shape: How shape of node-link diagrams impacts aesthetic appreciation and triggers interest," *i-Perception*, vol. 9, no. 5, 2018, Art. no. 2041669518796851.
- [129] U. Soni, Y. Lu, B. Hansen, H. C. Purchase, S. Kobourov, and R. Maciejewski, "The perception of graph properties in graph layouts," *Comput. Graph. Forum*, vol. 37, no. 3, pp. 169–181, Jun. 2018.
- [130] A. Lee and D. Archambault, "Communities found by users—Not algorithms: Comparing human and algorithmically generated communities," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2016, pp. 2396–2400.
- [131] K. Ballweg, M. Pohl, G. Wallner, and T. von Landesberger, "Visual similarity perception of directed acyclic graphs: A study on influencing factors and similarity judgment strategies," *J. Graph Algorithms Appl.*, vol. 22, no. 3, pp. 519–553, 2018.
- [132] M. Wattenberg, "Visual exploration of multivariate graphs," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2006, pp. 811–819.

- [133] H. Guo, J. Huang, and D. H. Laidlaw, "Representing uncertainty in graph edges: An evaluation of paired visual variables," *IEEE Trans. Vis. Comput. Graphics*, vol. 21, no. 10, pp. 1173–1186, Oct. 2015.
- [134] J. Schwank, S. Schöffel, J. Stärz, and A. Ebert, "Visualizing uncertainty of edge attributes in node-link diagrams," in *Proc. 20th Int. Conf. Inf. Vis. (IV)*, 2016, pp. 45–50.
- [135] J. Bae, T. Helldin, and M. Riveiro, "Understanding indirect causal relationships in node-link graphs," *Comput. Graph. Forum*, vol. 36, no. 3, pp. 411–421, 2017.
- [136] S. Schöffel, J. Schwank, and A. Ebert, "A user study on multivariate edge visualizations for graph-based visual analysis tasks," in *Proc. 20th Int. Conf. Inf. Vis. (IV)*, 2016, pp. 165–170.
- [137] W. Büschel, S. Vogt, and R. Dachsel, "Investigating link attributes of graph visualizations in mobile augmented reality," in *Proc. CHI Workshop Data Vis. Mobile Devices*, vol. 18, 2018, pp. 1–6.
- [138] W. Büschel, S. Vogt, and R. Dachsel, "Augmented reality graph visualizations," *IEEE Comput. Graph. Appl.*, vol. 39, no. 3, pp. 29–40, May 2019.
- [139] D. Yoon, N. H. Narayanan, S. Lee, and O. Kwon, "Exploring the effect of animation and progressive revealing on diagrammatic problem solving," in *Proc. Int. Conf. Diagrammatic Represent. Inference*, 2006, pp. 226–240.
- [140] R. M. Karim, O.-H. Kwon, C. Park, and K. Lee, "Study of colormaps in network visualization," *Appl. Sci.*, vol. 9, no. 20, 2019, Art. no. 4228.
- [141] W. Huang, P. Eades, and S.-H. Hong, "Measuring effectiveness of graph visualizations: A cognitive load perspective," *Inf. Vis.*, vol. 8, no. 3, pp. 139–152, Sep. 2009.
- [142] S. Ghani and N. Elmqvist, "Improving revisitation in graphs through static spatial features," in *Proc. Graph. Interface Conf.*, 2011, pp. 175–182.
- [143] C. Dunne and B. Shneiderman, "Motif simplification: Improving network visualization readability with fan, connector, and clique glyphs," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst.*, 2013, pp. 3247–3256.
- [144] B. Alper, N. H. Riche, and T. Höllerer, "Structuring the space: A study on enriching node-link diagrams with visual references," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2014, pp. 1825–1834.
- [145] B. Saket, P. Simonetto, S. Kobourov, and K. Börner, "Node, node-link, and node-link-group diagrams: An evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2231–2240, Dec. 2014.
- [146] B. Saket, C. Scheidegger, S. G. Kobourov, and K. Börner, "Map-based visualizations increase recall accuracy of data," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 441–450, Jun. 2015.
- [147] B. Saket, C. Scheidegger, and S. Kobourov, "Comparing node-link and node-link-group visualizations from an enjoyment perspective," *Comput. Graph. Forum*, vol. 35, no. 3, pp. 41–50, Jun. 2016.
- [148] R. Jianu, A. Rusu, Y. Hu, and D. Taggart, "How to display group information on node-link diagrams: An evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 11, pp. 1530–1541, Nov. 2014.
- [149] M. Okoe, S. S. Alam, and R. Jianu, "A gaze-enabled graph visualization to improve graph reading tasks," *Comput. Graph. Forum*, vol. 33, no. 3, pp. 251–260, Jun. 2014.
- [150] C. Ware and R. Bobrow, "Motion to support rapid interactive queries on node-link diagrams," *ACM Trans. Appl. Perception*, vol. 1, no. 1, pp. 3–18, Jul. 2004.
- [151] C. Ware, A. T. Gilman, and R. J. Bobrow, "Visual thinking with an interactive diagram," in *Proc. Int. Conf. Diagrammatic Represent. Inference*, 2008, pp. 118–126.
- [152] A. Skopik and C. Gutwin, "Improving revisitation in fisheye views with visit wear," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2005, pp. 771–780.
- [153] T. Moscovich, F. Chevalier, N. Henry, E. Pietriga, and J. Fekete, "Topology-aware navigation in large networks," in *Proc. Int. Conf. Hum. Factors Comput. Syst.*, 2009, pp. 2319–2328.
- [154] H. C. Purchase, B. Plimmer, R. Baker, and C. Pilcher, "Graph drawing aesthetics in user-sketched graph layouts," in *Proc. User Interface*, 2010, pp. 80–88.
- [155] H. C. Purchase, C. Pilcher, and B. Plimmer, "Graph drawing aesthetics—Created by users, not algorithms," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 1, pp. 81–92, Jan. 2012.
- [156] B. Plimmer, H. C. Purchase, and H. Y. Yang, "SketchNode: Intelligent sketching support and formal diagramming," in *Proc. 22nd Australas. Comput.-Hum. Interact. Conf.*, 2010, pp. 136–143.
- [157] T. Dwyer, B. Lee, D. Fisher, K. I. Quinn, P. Isenberg, G. Robertson, and C. North, "A comparison of user-generated and automatic graph layouts," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 961–968, Nov. 2009.
- [158] A. Drogemüller, A. Cunningham, J. Walsh, M. Cordeil, W. Ross, and B. Thomas, "Evaluating navigation techniques for 3D graph visualizations in virtual reality," in *Proc. Int. Symp. Big Data Vis. Immersive Anal. (BDVA)*, Oct. 2018, pp. 1–10.
- [159] S. Ghani, N. Elmqvist, and J. S. Yi, "Perception of animated node-link diagrams for dynamic graphs," *Comput. Graph. Forum*, vol. 31, no. 3, pp. 1205–1214, Jun. 2012.
- [160] D. Archambault and H. C. Purchase, "The mental map and memorability in dynamic graphs," in *Proc. IEEE Pacific Vis. Symp.*, Feb. 2012, pp. 89–96.
- [161] H. Lam, R. A. Rensink, and T. Munzner, "Effects of 2D geometric transformations on visual memory," in *Proc. 3rd Symp. Appl. Perception Graph. Vis.*, 2006, pp. 119–126.
- [162] K. Marriott, H. Purchase, M. Wybrow, and C. Goncu, "Memorability of visual features in network diagrams," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2477–2485, Dec. 2012.
- [163] B. Tversky, J. E. Corter, J. V. Nickerson, D. Zahner, and Y. J. Rho, "Transforming descriptions and diagrams to sketches in information system design," in *Proc. Int. Conf. Diagrammatic Represent. Inference*, 2008, pp. 242–256.
- [164] H. C. Purchase, "A healthy critical attitude: Revisiting the results of a graph drawing study," *J. Graph Algorithms Appl.*, vol. 18, no. 2, pp. 281–311, 2014.
- [165] L. Yu, J. V. Nickerson, and B. Tversky, "Discovering perceptions of personal social networks through diagrams," in *Proc. Conf. Diagrammatic Represent. Inference*, 2010, pp. 352–354.
- [166] X. Yuan, L. Che, Y. Hu, and X. Zhang, "Intelligent graph layout using many users' input," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2699–2708, Dec. 2012.
- [167] D. P. Singh, L. Lisle, T. M. Murali, and K. Luther, "Crowdlayout: Crowd-sourced design and evaluation of biological network visualizations," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. 1–14.
- [168] F. van Ham and B. Rogowitz, "Perceptual organization in user-generated graph layouts," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1333–1339, Nov. 2008.
- [169] C.-C. Lin, W. Huang, W.-Y. Liu, S. Tanizar, and S.-Y. Zhong, "Evaluating esthetics for user-sketched layouts of clustered graphs with known clustering information," *J. Vis. Lang. Comput.*, vol. 37, pp. 1–11, Dec. 2016.
- [170] C.-C. Lin, W. Huang, W.-Y. Liu, and W.-L. Chen, "Evaluating aesthetics for user-sketched layouts of symmetric graphs," *J. Vis. Lang. Comput.*, vol. 48, pp. 123–133, Oct. 2018.
- [171] C.-C. Lin, W. Huang, W.-Y. Liu, and C.-Y. Chen, "On aesthetics for user-sketched layouts of vertex-weighted graphs," *J. Vis.*, pp. 1–15, Sep. 2020. [Online]. Available: <https://link.springer.com/article/10.1007/s12650-020-00695-2>, doi: 10.1007/s12650-020-00695-2.
- [172] C. Fletcher, W. Huang, D. Arnese, and Q. V. Nguyen, "The role of working memory capacity in graph reading performance," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Apr. 2019, pp. 77–81.
- [173] W. Huang, "Evaluating graphs from a new perspective," in *Proc. IEEE 17th Int. Conf. Comput. Sci. Eng.*, Dec. 2014, pp. 1648–1652.
- [174] W. Huang, J. Luo, T. Bednarz, and H. Duh, "Making graph visualization a user-centered process," *J. Vis. Lang. Comput.*, vol. 48, pp. 1–8, Oct. 2018.



MICHAEL BURCH received the Ph.D. degree from the University of Trier, Germany, in 2010. He has worked for eight years as a Postdoctoral Researcher with the Visualization Research Center (VISUS), Stuttgart, and an Assistant Professor with the Visualization Group of Prof. Dr. Jarke J. van Wijk, Technical University of Eindhoven, The Netherlands. He is currently an Associate Professor with the Institute for Data Analytics, Visualization, and Simulation (DAVIS), University of Applied Sciences, Chur, Switzerland. His research interests include visual analytics, data science, information visualization, software visualization, graph visualization, and eye tracking. He has published more than 150 papers and articles in the field of visualization.



WEIDONG HUANG received the Ph.D. degree in computer science from the University of Sydney. He is currently an Associate Professor with the University of Technology Sydney, Australia. He also had formal training in experimental psychology and professional experience in psychometrics. He has more than 150 publications. His main research interests include human–computer interaction and visualization. He is also a founding Chair of the Technical

Committee on Visual Analytics and Communication for IEEE SMC Society and a Guest Editor of a number of SCI indexed journals. He has served as a Conference Chair, a PC Chair, or an Organization Chair for a number of international conferences and workshops.



MATHEW WAKEFIELD is currently pursuing the Ph.D. degree with the Swinburne University of Technology, Australia. He has a background in information technology, psychology, and social science. His current research interest includes applying connectionism for model-based machine learning.



HELEN C. PURCHASE received the Ph.D. degree from Cambridge University. She is currently a Senior Lecturer with the School of Computing Science, University of Glasgow, having previously held an academic position at The University of Queensland. She is engaged in research in two broad areas such as information visualization (in particular network visualization and empirical studies) and educational technology (in particular systems for collaborative learning and peer-

review). Her book *Experimental Human-Computer Interaction: A Practical Guide with Visual Examples* (Cambridge University Press, 2012), presents lessons learned and experimental guidance based on her several years of conducting visualization experiments—both successful and unsuccessful.



DANIEL WEISKOPF received the Dr. rer. nat. degree in physics from the University of Tübingen, Germany, in 2001, and the Habilitation degree in computer science from the University of Stuttgart, Germany, in 2005. He is currently a Professor with the University of Stuttgart. His research interests include information and scientific visualization, visual analytics, eye tracking, computer graphics, virtual and augmented reality, and special and general relativity.



JIE HUA received the B.S. degree in software engineering from Wuhan University, Wuhan, China, in 1999, and the Ph.D. degree in software engineering from the University of Technology Sydney, Sydney, Australia, in 2014. From 2016 to 2018 and in 2012, he was a Research Assistant with the Faculty of Engineering and Information Technology, University of Technology Sydney. In 2015, he has worked as a Research Assistant with the Business School, Western Sydney University, Sydney.

Besides, he has been working in the industry as the CTO and a Software Architect in Australia and China for more than 15 years. He is currently working as an A/Professor with the Faculty of Engineering and Information Technology, Shaoyang University, Shaoyang, Hunan, China. His research interests include software development, graph drawing, data visualization, big data, and deep learning.

...