Structuring the Space: A Study on Enriching Node-Link Diagrams with Visual References

Basak Alper
UC Santa Barbara
basakalper@gmail.com

Nathalie Henry Riche
Microsoft Research Redmond
nath@microsoft.com

Tobias Höllerer
UC Santa Barbara
holl@cs.ucsb.edu

ABSTRACT
Exploring large visualizations that do not fit in the screen raises orientation and navigation challenges. Structuring the space with additional visual references such as grids or contour lines provide spatial landmarks that may help viewers form a mental model of the space. However, previous studies report mixed results regarding their utility. While some evidence showed that grid and other visual embellishments improve memorability, experiments with contour lines suggest otherwise. In this work, we describe an evaluation framework to capture the impact of introducing visual references in node-link diagrams. We present the results of three controlled experiments that deepen our understanding on enriching large visualization spaces with visual structures. In particular, we provide the first tangible evidence that contour lines have significant benefits when navigating large node-link diagrams.

Author Keywords
Information Visualization, Network Diagrams, Navigation, Revisitation

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Exploring large visualization spaces that do not fit in the viewer's screen, such as a node-link diagram of several hundreds of nodes, raises orientation and navigation challenges [27]. Over the last decade, the human-computer interaction community has offered a plethora of interactive techniques to ease these tasks in large visualizations [8, 7, 29, 27, 23].

Enriching the space with additional visual elements such as grids [17] or contour lines [33] is an alternative approach. Such visual elements provide spatial landmarks which may help the viewer better grasp the dimensions of the space and provide memorable cues for better recognition and navigation [17] in large spaces.

Whether introducing non-data visual elements (such as a grid) hinders the readability of a visualization, and whether such elements play a beneficial role during an exploration are open-ended questions. While visualization practitioners strive to eliminate "chart junk" [34], prior work in psychology [12, 16] and empirical evidence presented by Bateman et al.[6] indicate that there may be benefits of even "seemingly useless" visuals in terms of memorability.

This paper aims to deepen our understanding on the utility of such visual enrichments within the context of visual representations of networks. Specifically, we study the effects of augmenting node-link diagrams with grid and contour lines. Both grids [5, 4] and contour lines [21, 13, 11] are commonly employed to enrich visualizations, however their relationship to the data differs significantly.

Grids are uniform structures independent of the data: two different datasets can be displayed on the exact same grid. In contrast, contour lines reflect an underlying property of the data such as a density measure, thus they present a unique pattern per dataset. Contour lines not only provide spatial landmarks enabling the use of spatial memory, but also reinforce data encoding by indexing elements topologically.

While a grid visualization has shown to be useful for navigation and revisitation of nodes in a node-link diagram [17], there is no conclusive evidence on the utility of contour lines. On the contrary, a controlled experiment by Tory et al.[33] on dot visualizations indicated that contours did not play a role on memorability.

Discrepancies in these results warrant further research. In this work, we describe an evaluation framework attempting to capture the impact of introducing additional visuals on the readability of node-link diagrams. In particular, we compare a uniform grid to contours generated from the data properties and investigate their potential role in helping the viewer build a mental model of the visualization space. Results of our controlled experiments shed a different light on previous results. We discuss the implications of our findings and reflect on key directions for further experimentations on the topic.

RELATED WORK
Grids are ubiquitous in a variety of charts that people use commonly such as scatterplots and bar charts. Their advocates argue that well-designed grids [5, 4] do not obstruct the primary visualization and serve as spatial references to compare relative positions of elements. For example, a grid in a scatterplot may help estimate the gap between two data points. However, we could not find any experimental results quantifying their benefits in visualizations.

Initial evidence by Ghani and Elmqvist [17] indicates that grids and other spatial landmarks may play a role in the

CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.

http://dx.doi.org/10.1145/2556288.2557112

formation of a mental model of large visualization spaces. Their study compared several designs for augmenting node-link diagrams, such as colored or textured grids and randomly distributed icons, while performing a revisitation task in a larger than screen node-link diagram. Their results show that solid colored grid cells increase performance when revisiting nodes. These findings suggest that grids play a role in the formation of a mental model of the space. However, the specific design used in this study contradicts suggested design guidelines [5, 4, 22]. The choice of employing grid cells with saturated solid colors may have improved revisitation performance at the cost of threatening the readability of the visualization.

Contour lines or topographic maps (contour lines with color filling and/or shading) are also commonly employed to enrich abstract data visualizations and node-link diagrams in particular [21, 13, 11]. Alluding to a geographic landscape featuring peaks and valleys, these representations enhance spatial metaphors. Fabrikant et al. argue that such "spatializations" act as fundamental sense makers for abstract domains because concepts about space are easily accessible to human cognition [15, 31]. However, they do not provide concrete evidence of spatial metaphors helping the viewer build a better mental model of the visualization space.

Tory et al. studied the benefits of 2D and 3D topographic maps for the readability of dot representations. Their first study dealt with the estimation of number of dots within a contour region [32]. They found that simple color coding was more effective than both conditions. A second study on memorability of dot representations [33], where subjects had to recognize previously seen dot distributions, also failed to show any benefit of 2D or 3D maps on memorability.

At the extreme end of the spectrum, Bateman et al.[6] have investigated the role of "chart junk", defined as seemingly not useful visual embellishments on the memorability of simple charts. These embellishments refer to illustrations or deformations that do not convey any information about the data, hence could be considered visual noise. The surprising results of their study suggest that such embellishments may provide richer spatial features that increase the memorability of charts. A recent study by Borkin et al.[10] support this argument by showing that visualizations with low *data-ink-ratio* and high visual density (more chartjunk) were more memorable than minimal, "clean" visualizations.

Other findings [9] on the topic suggest that there is a fine line between providing additional *data ink* (even if metaphorical) to increase memorability and engagement versus obstructing the data. These research pieces argue for conducting further studies in this area, carefully considering the design of visual embellishments and their impact on memorability and mental model.

In graph visualization domain, mental map preservation has been investigated in the context of dynamic graphs. However, work in this area focuses on the utility of layout preservation and animation rather than augmenting the representation with visual references [28, 1, 18].

Comparing and contrasting results across different experimental protocols, visual representations, tasks, datasets, and design of added visual enrichments proved extremely difficult. Conflicting outcomes of previous studies resorted in no clear answer whether or not augmenting a node-link diagram with additional visuals would help a viewer build a better mental model of the space. We refined a set of questions and propose a unified framework for evaluating the benefits and drawbacks of introducing enrichments to node-link diagrams.

EVALUATION FRAMEWORK

In this section we describe our rationale for assessing the role that visual structures such as grids and contours may play in the formation of a mental model. We seek answers to our high-level questions through low-level tasks performed in controlled settings.

Complementing visualizations with reference structures adds clutter that may hinder the readability of the primary visualization, and also contradicts guidelines of practitioners such as Tufte, advocating for *minimal ink*. Hence, assessing the potential negative impact of added visuals on readability of a visualization is crucial.

To address this question, we selected a low-level **readability task** on graph topology — *find common connections between two nodes*. We hypothesized that visual structures such as a grid and contours would interfere most with the links in node-link diagrams since their visual encodings are similar. Therefore, we selected an identify-common-neighbors task as outlined in [19], which requires participants to follow multiple links in the diagram.

A key reason for introducing visual references in a visualization is the idea that they can help the viewer form a better *mental model* of the space. Mental model is an overloaded term referring to concepts as diverse as the formation of an internal representation of a document collection and the formation of an internal 3D model of an object based on its 2D representation. In this paper, our definition of mental model is *an internal representation that one constructs about the spatial organization of objects*.

We believe that forming a mental model of the space involves at least two aspects of the spatial organization of visual elements: (1) *spatial structure*: gaining a sense of the relative positions of the elements in the visualization (i. e. grasping their distance and orientation to each other); (2) *spatial landmarks*: memorizing unique recognizable motifs in the representation that the viewer can refer to in order to navigate and orient himself in the space.

To measure the role of grids or contour lines in helping the viewer gain a sense of spatial structure and relative positioning of elements, we selected a **comparison task** — *determine if and how two graphs differ*. We conjectured that participants would make use of relative spatial encodings (e.g. the cluster to the left of the central grid cell) when performing a region-by-region comparison between two similar visualizations displayed alongside each other.

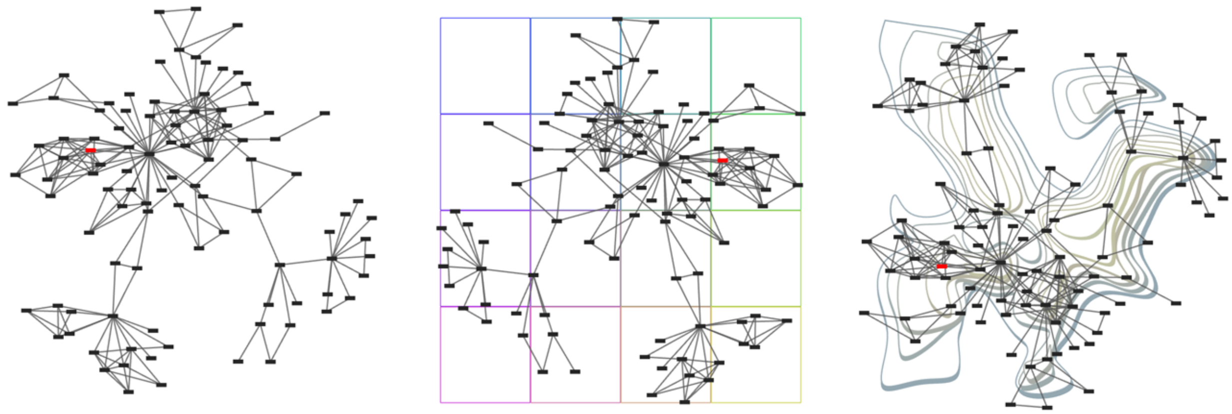

Figure 1. Example stimuli images from our controlled study evaluating user performance in revisiting previously highlighted nodes. Shown are the three visualization conditions used in the experiment: node-link diagram only, node-link diagram with grid, and node-link diagram with contour maps.

Gaining an understanding of recognizable and memorizable motifs in node-link diagrams is a research question in itself [26]. While a recall task such as drawing the diagram from memory may provide useful insights [26], it is not applicable for graphs with more than ten nodes. To measure the role of grids and contour lines in providing memorable spatial features, we devised a **revisitation task** — *revisit previously highlighted nodes*. We selected this task to build upon the previous work of Ghani et al.[17], recognizing it also as a plausible task that analysts perform [24, 30].

We also hypothesized that the construction of a mental model differs significantly depending on whether the whole visualization space can be seen in one view or not. While in the former case visual memory plays a more dominant role, spatial memory might have a stronger effect in the latter. For this reason, we suggest to perform the revisitation task under two different interaction conditions. We propose a first condition in which the viewer can only pan in the visualization space, thus never seeing the whole visualization space at once; and a second condition in which the viewer can acquire a view of the whole space with a zoom-out operation.

EVALUATION

We performed three studies for each task described above.

Techniques

We compared grid and contour lines and with a control condition of plain node-link diagram. Technique settings and data were adjusted for each of the separate experiments and are reported in each study section.

Node-Link-Only.

In this control condition, participants complete the tasks using a node-link diagram without any additional visuals.

Grid.

This condition augments the node-link diagram used in the previous condition with an underlying grid. We employ a modified version of the grid design proposed by Ghani et al.[17] which used strongly saturated solid colors for each grid cell. In realistic scenarios however, such colors may greatly interfere with the visual encoding of the primary data [35, 4], a disadvantage we believe overpowering the benefit of improved memorability in many visual analysis tasks. Thus, we opted for a grid representation consisting of empty frames rather than solid color regions. We assigned a distinct color to each grid cell as in [17], but used transparency as advised in [4].

Contours.

This condition augments the node-link diagram used in the control condition with contour lines. In the context of node-link diagrams, contour lines can be derived from several data properties such as node degree or node centrality. Similar to [21] we generate our contours starting with Delaunay triangulation of nodes in a given layout. Corners of each triangle are assigned a height value corresponding to the selected node property. We interpolate this height value within each triangle using barycentric coordinates. Given a set of iso-property values indicating where the contours should pass, our implementation of the marching squares [25] algorithm computes the set of points marking the boundary of the iso-value. We sample these boundary points and use piecewise Bezier curves to connect them.

Visual references such as grids and contours are designed to reside in a visual middle ground between the foreground of the primary data encodings and the background, such that they can be brought to the foreground when attended to, and pushed to the background otherwise [20, 4, 5]. Following the design guidelines suggested in [4, 5], we chose visual parameters to render the contours sufficiently discernible, yet minimize their interference with the primary visualization. Instead of a gray-scale representation, we opted for color coding to help distinguish nested contour lines. However, we selected a very narrow color palette, with minimal saturation and high luminance, allowing it to blend more with the white background of the node-link diagram. In addition, we utilized shadow-like offsets to emphasize unique contour shapes further, while making them more distinguishable from the simple stroke used for the links of the node-link diagram. Another motivation was to allude to a 3D landscape, featuring recognizable geographic formations such as peaks and valleys.

Experiment 1: Readability

This experiment investigates whether the presence of grid or contour lines compromises the readability of a node-link diagram by employing a common neighbors task. We performed a within-subject design with 3 Techniques $\times$ 2 Data Sizes $\times$ 4 Repeats. We recruited 9 participants, 3 females and 6 males, with a mean age of 33.8 years.

Data

For this study, we generated graphs using Eppstein and Wang's power law model [14]. The model is designed to ensure that the synthetic data generated follows a well-known property of the degree of node observed in many real networks [3]. We used two different sizes: 50 nodes and 75 links (small); and 100 nodes and 150 links (large). We generated 6 graphs per size for each repeat and training trials. Across techniques, we used the same graphs. In order to minimize learning effects, we flipped the layout horizontally and/or vertically for each technique.

Technique Settings

All graphs, regardless of their size, were scaled to fit in an area of 600 by 600 pixels, fitting on a standard computer screen and removing the need for interactive navigation that could impact the completion time. We opted for a legible but small visualization size to capture the potential effect of background interference with a high density of visual elements. From a series of pilots, we found that a grid of 3 by 3 provided a good compromise. For the contour lines, we experimented with several metrics and opted for using the node degree. We observed that this metric produced contours distinctive enough from each other, while not introducing too much clutter. We used the same color palette for coloring the contours and the grid lines. Figure 2 shows examples from visual stimuli used in our experiment.

Procedure

Given two highlighted nodes, participants had to select nodes that are connected to both of the highlighted nodes in a total of 24 trials. The order of techniques were counterbalanced across participants. For each trial we selected two random high-degree nodes which were not connected directly but had a minimum of two common neighbors. Due to the high connectivity requirements, the highlighted nodes were likely to appear at a busy region of the node-link diagram, where the visualization is more susceptible to interference with the background. Thus, the chance of capturing a potential effect of the background on the readability of the node-link diagram would be higher. The order of techniques were counterbalanced across participants. The participants selected the common neighbors by clicking on the nodes. After selecting all common neighbors, the participants were asked to confirm their answer by pressing a button. Participants were instructed to take breaks as needed after every trial while the visualization screen was blank.

Hypothesis

We hypothesized that **(H1)** Grid and Contours would decrease the readability of the representation and impact at least the completion time as they introduce additional visual clutter.

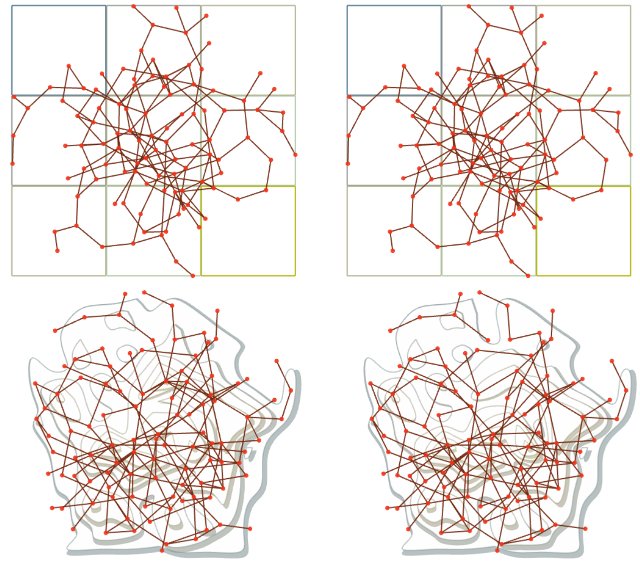

Figure 2. Examples of the visual stimuli used in the controlled studies on readability (Experiment 1) and comparison (Experiment 2). Only one of each pair of images was used in the readability study.

Results

Since accuracy results did not follow a normal distribution, we analyzed them using Friedman's non-parametric test. Friedman's test did not reveal any significant difference in accuracy between Node-Link-Only, Grid, or Contours. All techniques had a mean of about 90% accuracy (with 3.5% standard error). Pairwise comparisons of Wilcoxon's test Z values reveal weak effect sizes ($r < 0.1$).

We analyzed the completion time for correct trials only, using a Mixed Linear Model (MLM) that can handle missing values. We excluded the about 10% of incorrect trials. MLM did not reveal any significant difference in completion time, contradicting our hypothesis **(H1)**. All techniques had a mean of about 15 seconds (with 7 sec standard error).

We asked the participants to rate whether they found the Grid or Contours distracting for this task, using a 5-point Likert scale from 1 (not distracting at all) to 5 (very distracting). Grid received an average rating of 1.5 (SD=0.35), while the average rating for Contours was 2.3 (SD=0.47). In line with **(H1)**, the majority of participants (6 out of 9) preferred the Node-Link-Only condition for this task.

Experiment 2: Visual Reference

In this experiment participants were asked to identify presence and nature of differences between two seemingly similar graphs that are displayed side by side. We performed a within-subject experiment of 3 Techniques $\times$ 2 Data Sizes $\times$ 4 Repeats. This experiment was performed right after the Experiment 1 session by the same participants using the same data sets and visual stimuli.

Procedure

For each graph, we generated an altered version by randomly deleting a high degree node with all its connections. However, we avoided eliminating nodes with high betweenness centrality to ensure that the task was challenging enough. We created new contour lines for the altered graph since their node degree had changed. Half of the trials presented the exact same graph on both sides. Participants had to answer whether the graph on the right had a smaller, higher, or equal number of nodes and edges as the graph on the left by pressing one of the three buttons.

Hypotheses

We hypothesized that **(H2)** participants would be more effective to conclude if the two graphs are identical with Contours, where they can rely on these visuals for gauging the differences instead of carefully inspecting the node-link diagram. For quantifying the differences (larger or smaller), we expected that **(H3a)** Grid would yield better results than Node-Link-Only, as grid cells would enable participants to structure the space and inspect regions systematically. We also believed that **(H3b)** Contours would yield to better results than the two other techniques because the landscape built from data properties would guide participants' attention towards the region that is different.

Results

Since accuracy results did not follow a normal distribution, we analyzed them using Friedman's non-parametric test. The test revealed a significant difference between techniques for identifying if graphs are identical or not ($p < .05$). Pairwise comparisons using Wilcoxon's tests indicate that the Contours technique outperforms Node-Link-Only ($p < .05$) and Grid ($p < .01$) confirming **(H2)**. An estimate of effect size based on Z value reveals a medium effect size ($r \approx .3$) for both of these pairwise comparisons. Mean accuracy indicates that Contours are 10% more accurate than Node-Link-Only and 18% more accurate than Grid. Graph size did not yield to any significant difference between techniques.

We analyzed the results on completion time for correct trials only using a Mixed Linear Model (MLM) that can handle missing values. We excluded the 24% of incorrect trials (*i.e.* participants failed to recognize if the modified graph was identical, smaller or larger). MLM revealed an effect of Technique ($F_{2,16} = 5.29, p < .05$) and Size ($F_{1,8} = 41.49, p < .001$) on completion time. As expected, pairwise comparison revealed that participants performed 20% faster to identify changes in Small size networks, with a statistically significant difference ($p < .001$). Pairwise comparisons revealed that Contours led to significantly faster task completion times ($p < .05$) than the other two techniques. Mean completion time indicates that users performed the tasks 12% faster on average with Contours than with the two other techniques, confirming **(H3b)**. Contrary to **(H3a)**, Grid however did not produce more accurate or faster identification of differences between graphs compared to Node-Link-Only.

We asked participants to rate the usefulness of the Grid and Contours visualizations for this comparison task, using a 5-point Likert scale from 1 (not useful at all) to 5 (very useful). Grid received an average rating of 3.77 (SD=0.22), while the average rating for Contours was 4.45 (SD=0.24). 6 out of 9 participants preferred the Contours visualization, while the remaining 3 preferred the Grid visualization for this task.

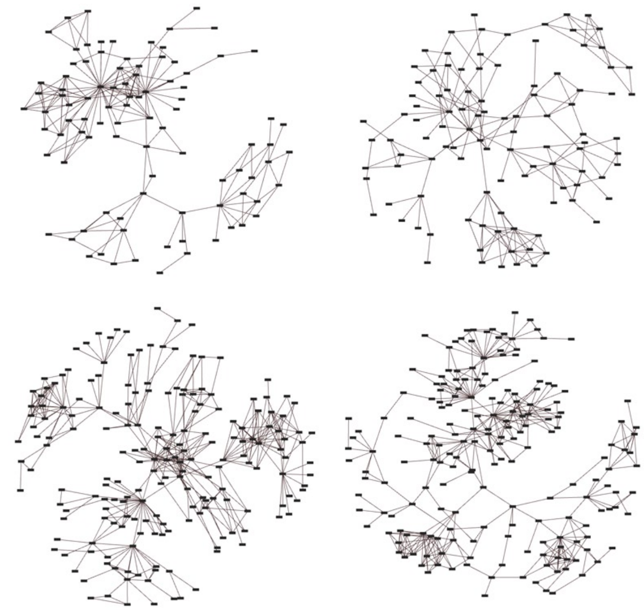

Figure 3. Examples from the small (top row) and large (bottom row) data sets used in Experiment 3.

Experiment 3: Revisitation

In this experiment, we investigated whether the Grid and the Contours visual reference structures aided users in forming a mental model of a visualization space. To assess the formulation of a mental model, we devised a task of revisiting previously highlighted nodes in a visualization space that is larger than the viewing window. We performed a 3 Techniques $\times$ 2 Data Sizes $\times$ 2 Interaction Methods $\times$ 3 Repeats within-subject design. We recruited 12 participants, 8 males and 4 females, with a mean age of 28.3 years.

Data

We used two data sizes in this experiment : 100 nodes and 150 links (small), and 150 nodes and 300 links (large). We decided against using a graph generator for this task because we did not have access to one that could produce graphs with distinguishable topological features (i. e. dense communities or hubs) as real graph often do [2]. We conjectured that this factor could bias our results as grid or contour lines are likely to play a more significant role when used to augment "featureless" visual spaces. Therefore, we generated our graphs from real social network data. Starting from a data set representing co-autorship relationships in a real social network, we randomly added and removed nodes and edges until the desired number of elements and sufficient distinctive topological features were achieved. We generated five graphs per size for each repeat and training trials. Examples are shown in Figure 3. Across techniques, we used the same graphs. In order to minimize learning effects, we flipped the layout horizontally and/or vertically for each visualization technique.

Technique Settings

In this experiment, participants had to perform the task in a visualization space that did not fit in view (unless specifically zoomed out in one of our conditions). We used a window size of 1000x1000 pixels over a 3000x3000 pixels canvas (small)

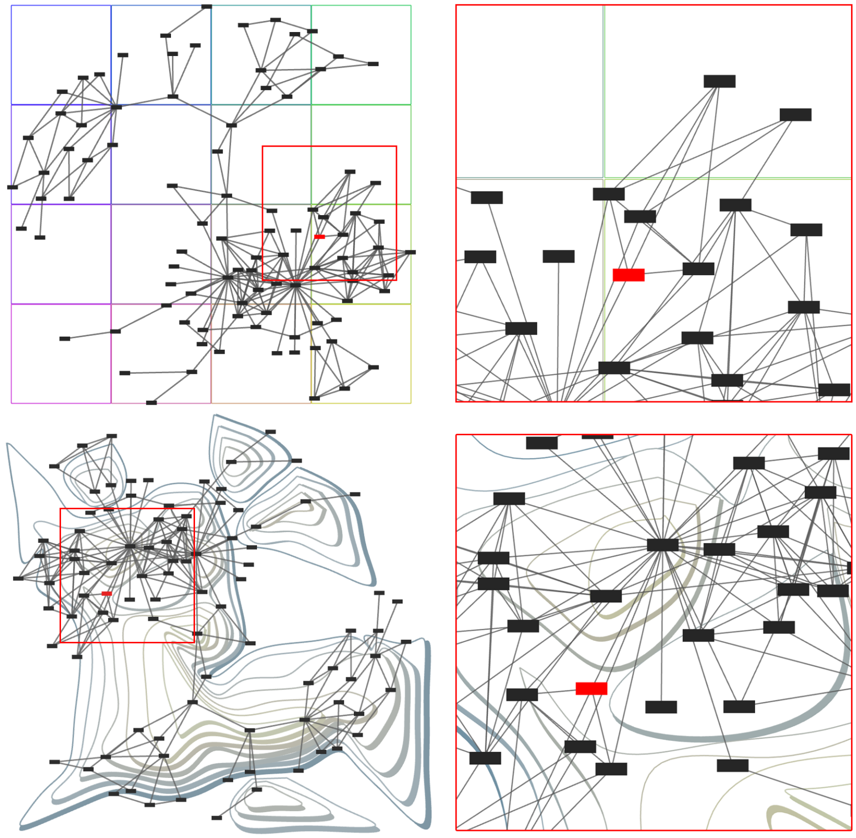

Figure 4. Examples from the Grid and Contour visualization stimuli images used in Experiment 3 (revisitation). The images on the left show the whole visualization space with a highlight of what can fit in the view in the pan and zoom-in conditions. The images on the right show the contents of the viewing window in detail.

or a 4500x4500 pixels canvas (large). We picked a grid cell size of 750x750 pixels ensuring that an entire cell and some portion of its neighboring cells would fit in the view, while avoiding a fine grid spacing. Thus, we used a 4x4 grid for the small data, and a 6x6 grid on the large one. We opted for using color on the grid lines as a secondary cue to encode relative positions of the grid cells (see Figure 4, top).

To generate contour lines, we used the betweenness centrality of each node. This metric produced a good compromise between the simplicity of the patterns (i.e. the number of distinct peaks) and the presence of distinguishable features (i.e. the variation in shape). See Figure 4, bottom for an example visual stimulus.

Navigation and Interaction

The participants used either pan or zoom to navigate within the visualization space. Dissociating navigation techniques is common practice [23] as it ensures similar navigation strategies between participants. Here, it also enabled us to observe different strategies while forming a mental model of the space. In the pan interaction mode, participants could not view the whole visualization but they could pan by dragging their mouse. In the zoom interaction mode, we designed a simple two-level zoom wherein participants could toggle between a zoomed-in and a zoomed-out state on right mouse button click as implemented in [23]. In the zoomed-out state, the whole visualization fit in the view. In this mode, when participants right-clicked, the visualization zoomed in, centering around the position of the mouse click. The scales used in the zoomed-in state and the pan mode were the same.

Procedure

We designed a revisitation task consisting of a learning and an execution phase, similar to [17]. In the learning phase, three random nodes were highlighted in sequence. The starting view was at the center of the visualization space in the pan mode, and at the zoomed-in state for the zoom mode. Participants had to navigate to find the highlighted node, collect information about its locality and click on it to view the next highlighted node. Once they clicked all three nodes, they viewed a blank screen for four seconds. After, the execution phase began with the starting view, where participants had to revisit the nodes (now dehighlighted) in the same order. They provided their answer by clicking on the node. Time to revisit a node was limited to 25 seconds, with a countdown appearing in the last 10 seconds. If the time expired, a pop up dialogue instructed them to proceed to the next node. We forced a ten second break after every trial, displaying a photograph to reset visual memory. We instructed them to take longer breaks if needed.

Participants completed 18 trials using pan, and another 18 using zoom interaction. We used the same visuals but highlighted a different set of nodes to limit learning effects. We counterbalanced the order of visualizations, but we kept the order of interaction methods fixed across participants, opting for presenting the most difficult condition (pan) first. We trained participants before each task and interaction mode, informing them about strategies they could utilize. We instructed them to collect sufficient information about the locality of a highlighted node before clicking on it in the learning phase while being as time efficient as possible. On average participants took about 70 minutes to complete the study. All participants were rewarded with a 25 dollar gift card, while the best performing participant was promised an additional 25 dollars.

Hypotheses

We expected that **(H4a)** participants would be least accurate with the Node-Link-Only condition, because the presence of Grid or Contours would provide additional visual references encoding the locality of a visited node. In addition, we expected **(H4b)** Contours to outperform Grid because, in contrast to the uniform visual characteristics of grids, contours provide unique visual structures in different parts of the visualization which would ease recall.

We hypothesized that **(H5)** in the zoom condition, differences between the Contours and Grid techniques may not be significant since participants could effectively memorize the exact row and column of the cell containing the node in the Grid visualization. However, we expected that **(H6)** in the pan condition, Contours would have clearer benefits over the other two techniques as they provide unique landmarks across space.

Finally, we did not expect **(H7)** to see any significant time difference across visualization techniques because we hypothesized that any possible gains provided by the visuals on recalling the location would be eclipsed by the interaction time required to navigate there.

Results

Since accuracy did not follow a normal distribution, we analyzed them using Friedman's non-parametric test, which revealed a significant difference among techniques ($p < .001$). Pairwise comparisons using Wilcoxon's tests showed that Contours outperformed Node-Link-Only ($p < .01$) and Grid ($p < .001$), confirming **(H4b)**. An estimate of effect size based on Wilcoxon's test Z value revealed medium effect size ($r \approx .3$) between Contours and the other two techniques, while the effect size between Node-Link-Only and Grid was small ($r < .1$). Mean accuracy indicates that Contours produced 7% more accurate results than Node-Link-Only and 16% more accurate results than Grid for revisiting previous nodes. Contrary to our hypothesis **(H4a)**, we did not find any significant difference between Grid and Node-Link-Only. When splitting results by interaction condition, Friedman's tests revealed significant differences in accuracy among techniques in both the zoom ($p < .05$) and pan ($p < .01$) conditions. Contours significantly outperformed Node-Link-Only ($p < .05$) and Grid ($p < .05$) using zoom, contradicting **(H5)**. In addition, Contours outperformed Grid ($p < .001$) but not Node-Link-Only in the pan condition, partially validating **(H6)**. Table 1 and Figure 5 summary the results.

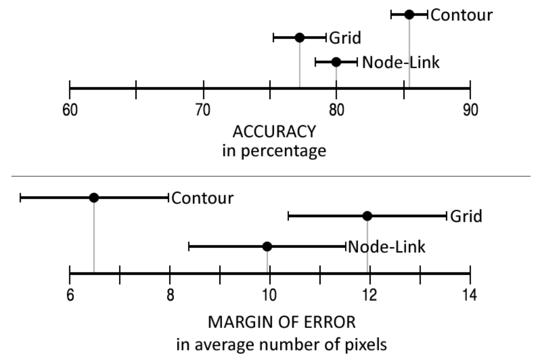

Figure 5. Accuracy and margin of error results for all three techniques in Experiment 3. Means and +/- standard error values are shown.

Technique	Pan and Zoom	Pan	Zoom
C x G x NL	($p < .001$)	($p < .01$)	($p < .05$)
C x G	($p < .001$)	($p < .001$)	($p < .05$)
C x NL	($p < .01$)		($p < .05$)

Table 1. Significant differences in accuracy across techniques and interaction conditions in Experiment 3. C: Contour, G: Grid, NL: Node-Link-Only

We analyzed the results on completion time for correct trials only using a Mixed Linear Model (MLM) that can handle missing values. We excluded 24% of incorrect trials. Inline with **(H7)**, MLM did not reveal any significant difference among techniques.

Subjective Ratings

After the controlled experiment, participants rated their confidence in their results using a 5-point Likert scale from 1 (not confident) to 5 (very confident). Mean confidence levels are provided in Table 2 and Figure 6. Since these ratings do not follow a normal distribution, we analyzed them using non-parametric tests. Friedman's test shows a significant difference in participant confidence between techniques for both pan ($p < .05$) and zoom ($p < .001$). Pairwise comparisons using Wilcoxon's test reveal that participants felt significantly more confident using the Contour compared to Node-Link-Only for both pan ($p < .05$) and zoom ($p < .01$). In the zoom condition, participants felt also significantly more confident using Contour than using Grid ($p < .05$).

We also asked the participants to give an overall preference ranking for the three visualization techniques. Friedman's test indicated a significant difference in preference among techniques ($p < .01$). Pairwise comparisons using Wilcoxon's test reveal that Contours were ranked significantly higher than both Node-Link-Only ($p < .001$) and Grid ($p < .05$). Indeed, 8 out of 12 participants ranked Contour as their favorite technique.

Finally, we also report on most recurrent participants' comments collected using a printed questionnaire. For Node-Link-Only, 5 out of 12 participants indicated that they felt their performance was dependent on the node to revisit and particularly easy for nodes on the periphery of the diagram. For the Grid, 5 out of 12 participants indicated that they were able to relocate the grid cell but that they forgot the exact position of the node within this cell. A few participants explained

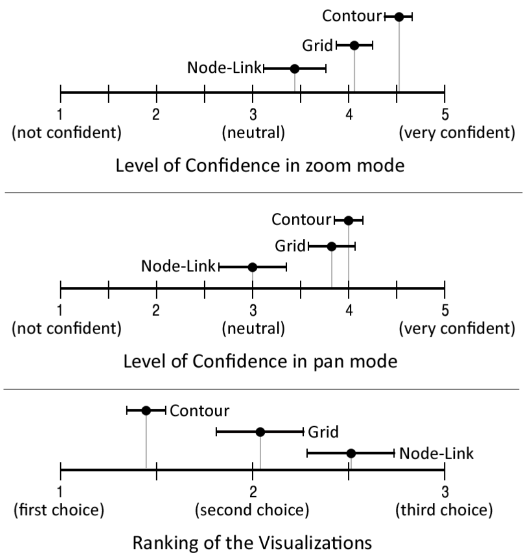

Figure 6. Subjective rating results for confidence levels in the zoom (top), and pan interaction modes (middle), as well as the overall rankings of the three visualization techniques (bottom). Means and +/- standard error values are shown.

Int.	Contours	Grid	Node-Link-Only
Pan	4 (SD=0.6)	3.83 (SD=1.02)	3 (SD=1.27)
Zoom	4.58 (SD=0.51)	4.16 (SD=0.57)	3.41 (SD=1.16)

Table 2. Average confidence levels indicated by the participants for each interaction mode per visualization technique. Standard deviations are provided in parenthesis.

that counting the grid cell was diverting their attention from the actual graph structure. For the Contour, 7 out of 12 participants indicated that they used contour shapes as spatial landmarks. For example, a participant commented that he "could see islands, coves and mountains to easily associate with geographic locations" and another explained that he "could find unique representative shapes... that represent objects in real life, for example, crown on a prince's head".

DISCUSSION

In this section we discuss the implications of our findings as well as address the key limitations of our studies to take into account in the future.

Implications for Design

Whispering Visuals

We followed the guidelines prescribed in [5] for the visual design of our grids and contour lines. Our readability study did not reveal any significant difference when additional visual structures were present. As the common neighbor task required participants to follow many links in the diagram that were overlapping these visuals, we feel confident that readability of the diagram was not hindered when suggested design guidelines were followed.

Structuring the Space with Grids

We were surprised that our results did not show any significant increase in performance using Grid versus Node-Link-Only. In particular, the results from our revisitation study seem to be in contradiction with previous studies reported in [17]. We explain these discrepancies by two major differences in the experimental setups. First, in contrast to the previous study, we selected a more "whispering" color palette. This choice may have downplayed the role of grids in helping user form a mental model of the space but ensured its readability as advised in [5, 4] and confirmed in our readability study. Second, we selected graph data featuring distinguishable topological structures, as we conjectured it was more common in real datasets. However, such features may have helped task performance more than the information provided by the grid. Experiments with representations that naturally contain less salient motifs (such as dot visualizations) may capture a stronger effect of a grid.

Unexpectedly, despite the lack of a significant increase in performance, the grid was consistently preferred over the control condition. Participants also reported a higher sense of confidence in their answers with the grid. These subjective impressions could point to benefits that we failed to capture with our current experimental settings. However, from our current findings, we also want to call for a careful consideration of their use in visualizations, as they might lead to a false sense of confidence.

Structuring the Space with Contour Lines

While many researchers experimented with augmenting visualizations with contour lines or topographic maps [33, 33, 15, 21, 13, 11], there has been no tangible evidence of their utility. On the contrary, advocates for *minimal ink* may caution against the use of these structures as they may obstruct readability. Our quantitative findings are the first to indicate that contour lines do not hinder readability (study 1), play a significant role in helping viewers structure the space (study 2), and form a mental model when navigating a large visualization (study 3). Our quantitative results from the controlled experiments as well as qualitative comments from the participants suggest that contour lines can provide a general outlook of the node-link diagram, helping viewers identify unique spatial landmarks to return to previously visited nodes. Our subjective ratings also concur with the increased performance as contour lines are consistently preferred. Based on these insights, we project that contour lines could ease navigation and orientation in an infinite zoomable and pannable canvas.

In our studies we manually selected the metric to generate the contours to ensure the presence of identifiable geometric features while minimizing interference with the primary visualization. While node centrality provided sufficient discernible characteristics for a social network type of data (study 3), we opted to use the degree metric for synthetic data generated using a power law distribution (study 1, 2). Non-graph data properties associated with the nodes can also be used. Thus, the metric to be used for contour generation is highly specific to the data.

Structuring the Space for Dynamic Visualizations

Enriching node-link diagrams with visual structures that reflect a data property is a double-edged sword. It may, as we illustrated in studies 2 and 3, help viewer's build a mental model of the representation. However, in contrast with a grid, these structures may also prove highly unstable when exploring a dynamically changing diagram. Contour lines could help grasp where the changes are, but if these are dramatic, they may rather disorient the viewer.

Study Limitations

In our studies we strived to devise realistic tasks while controlling many factors that could impact our findings. Consequently, as with all controlled experiments, there are key limitations that need to be taken into account when generalizing our results and designing further experiments.

Tasks and Sample Size

What a mental model exactly is and how to assess it is an open and challenging research question. Our evaluation framework is an attempt at characterizing several of the aspects that may be at play when viewers form a mental model of a visualization. However, other aspects could be taken into account. In particular, we did not include any recall [26] or recognition tasks [33] as we could not decide on a systematic way to operationalize them or interpret their results.

Our study is of a limited sample size. For the first two studies, where we recruited only nine subjects, arriving at definite conclusions is not possible. However, we can observe certain trends. For instance, due to very small effect size, only a much larger population could reveal a significant effect of visual enrichments on readability. Our study on identifying differences across similar graphs unexpectedly showed that grid visualizations might not be helpful for side-by-side comparison tasks. This trend calls for further investigations on the role of a grid as a reference cue.

Granularity of the Visual Enrichments

As mentioned earlier, we manually adjusted granularity settings such as the number of contour lines and grid cells after a series of pilot sessions, ensuring that a certain amount of contour or grid lines were always present, regardless of the scale. We still collected conflicting feedback from our participants: 3 commented that smaller cells would allow them to better structure the space; 2 argued for larger cells as they had difficulties remembering the cell positions. Participants also commented that while some contour lines did not appear to provide enough recognizable patterns in the pan and zoom-in modes, they became discernible as spatial landmarks in the zoom-out mode. While we had to fix granularity of both contours and grids for a manageable study size, future studies that focus on granularity as an independent variable are needed.

Semantics

In this study we did not address the particular role of the data semantics. However, we would like to acknowledge that semantics may play a major role in mental model formation, possibly eclipsing the gain provided by additional visual structures. However, this role would be highly dependent on the dataset itself and the participants' previous knowledge, interests and analysis skills, thus making this factor extremely difficult to control. We believe that studies in less controlled settings (ideally longitudinal ones where participants explore their own datasets) would be more adequate to evaluate the interplay between semantics and additional visual structures when a viewer is forming a mental model of the visualization.

CONCLUSION

In this paper, we proposed an evaluation framework to extend our knowledge on the impact of introducing visual references to node-link diagrams. Through three controlled studies, we collected evidence quantifying the benefits and drawbacks of two types of visual structures : grids and contour lines.

We discussed how our findings shed a new light on previous results in the literature. In particular, our results seem to indicate that when node-link diagrams present recognizable motifs, such as clusters, the benefits of grids for revisitation tasks [17] may not be discernible. Our studies are also the first ones to reveal tangible benefits of the use of contour lines to enrich node-link diagrams. Our experimental design attempts to capture how these visuals could help the viewer form a mental model of the visualization (by structuring the space and identifying spatial landmarks). Results indicate that contour lines play a significant role in these activities.

We believe that our results on the benefits of enriching visualizations with additional visual structures will inspire further research in this space, and the evaluation framework we propose will serve as a building block for further studies.

REFERENCES

1. Archambault, D., Purchase, H., and Pinaud, B. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *Visualization and Computer Graphics, IEEE Transactions on 17*, 4 (2011), 539–552.

2. Bach, B., Spritzer, A., Lutton, E., Fekete, J.-D., et al. Interactive random graph generation with evolutionary algorithms. In *Graph Drawing* (2012).

3. Barabási, A.-L., and Albert, R. Emergence of scaling in random networks. *science 286*, 5439 (1999), 509–512.

4. Bartram, L., Cheung, B., and Stone, M. C. The effect of colour and transparency on the perception of overlaid grids. *Visualization and Computer Graphics, IEEE Transactions on 17*, 12 (2011), 1942–1948.

5. Bartram, L., and Stone, M. C. Whisper, don't scream: Grids and transparency. *Visualization and Computer Graphics, IEEE Transactions on 17*, 10 (2011), 1444–1458.

6. Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., and Brooks, C. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the 28th international conference on Human factors in computing systems*, ACM (2010), 2573–2582.

7. Baudisch, P., and Rosenholtz, R. Halo: a technique for visualizing off-screen objects. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2003), 481–488.

8. Bederson, B. B., and Hollan, J. D. Pad++: a zooming graphical interface for exploring alternate interface physics. In *Proceedings of the 7th annual ACM symposium on UIST*, ACM (1994), 17–26.
9. Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P. W., Reppa, I., Floridi, L., and Chen, M. An empirical study on using visual embellishments in visualization. *Visualization and Computer Graphics, IEEE Transactions on 18*, 12 (2012), 2759–2768.
10. Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., and Pfister, H. What makes a visualization memorable? *Visualization and Computer Graphics, IEEE Transactions on 19*, 12 (2013), 2306–2315.
11. Cao, N., Sun, J., Lin, Y.-R., Gotz, D., Liu, S., and Qu, H. Facetatlas: Multifaceted visualization for rich text corpora. *Visualization and Computer Graphics, IEEE Transactions on 16*, 6 (2010), 1172–1181.
12. Collins, A. M., and Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological review 82*, 6 (1975), 407.
13. Cortese, P. F., Di Battista, G., Moneta, A., Patrignani, M., and Pizzonia, M. Topographic visualization of prefix propagation in the internet. *Visualization and Computer Graphics, IEEE Transactions on 12*, 5 (2006), 725–732.
14. Eppstein, D., and Wang, J. A steady state model for graph power laws. *arXiv preprint cs/0204001* (2002).
15. Fabrikant, S. I., and Skupin, A. *Cognitively plausible information visualization*. Elsevier, 2005.
16. Gambrell, L. B., and Jawitz, P. B. Mental imagery, text illustrations, and children's story comprehension and recall. *Reading Research Quarterly* (1993), 265–276.
17. Ghani, S., and Elmqvist, N. Improving revisitation in graphs through static spatial features. In *Proceedings of Graphics Interface 2011*, Canadian Human-Computer Communications Society (2011), 175–182.
18. Ghani, S., Elmqvist, N., and Yi, J. S. Perception of animated node-link diagrams for dynamic graphs. In *Computer Graphics Forum*, vol. 31, Wiley Online Library (2012), 1205–1214.
19. Ghoniem, M., Fekete, J.-D., and Castagliola, P. A comparison of the readability of graphs using node-link and matrix-based representations. In *Information Visualization, 2004.*, IEEE (2004), 17–24.
20. Gombrich, E. H., Gombrich, E. H., Gombrich, E. H., and Gombrich, E. H. *Art and illusion: A study in the psychology of pictorial representation*, vol. 5. Phaidon London, 1977.
21. Gronemann, M., and Jünger, M. Drawing clustered graphs as topographic maps. In *Graph Drawing*, Springer (2013), 426–438.
22. Heer, J., and Bostock, M. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the 28th ACM SIGCHI*, ACM (2010), 203–212.
23. Irani, P., Gutwin, C., and Yang, X. D. Improving selection of off-screen targets with hopping. In *Proceedings of the ACM SIGCHI conference*, ACM (2006), 299–308.
24. Lee, B., Plaisant, C., Parr, C. S., Fekete, J.-D., and Henry, N. Task taxonomy for graph visualization. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, ACM (2006), 1–5.
25. Lorensen, W. E., and Cline, H. E. Marching cubes: A high resolution 3d surface construction algorithm. In *ACM Siggraph Computer Graphics*, vol. 21, ACM (1987), 163–169.
26. Marriott, K., Purchase, H., Wybrow, M., and Goncu, C. Memorability of visual features in network diagrams. *Visualization and Computer Graphics, IEEE Transactions on 18*, 12 (2012), 2477–2485.
27. Moscovich, T., Chevalier, F., Henry, N., Pietriga, E., and Fekete, J.-D. Topology-aware navigation in large networks. In *Proceedings of the 27th international conference on Human factors in computing systems*, ACM (2009), 2319–2328.
28. Purchase, H., Hoggan, E., and Görg, C. How important is the "mental map"?–an empirical investigation of a dynamic graph layout algorithm. In *Graph drawing*, Springer (2007), 184–195.
29. Shoemaker, G., and Gutwin, C. Supporting multi-point interaction in visual workspaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 999–1008.
30. Skopik, A., and Gutwin, C. Improving revisitation in fisheye views with visit wear. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2005), 771–780.
31. Skupin, A., and Fabrikant, S. I. Spatialization methods: a cartographic research agenda for non-geographic information visualization. *Cartography and Geographic Information Science 30*, 2 (2003), 99–119.
32. Tory, M., Sprague, D. W., Wu, F., So, W. Y., and Munzner, T. Spatialization design: Comparing points and landscapes. *Visualization and Computer Graphics, IEEE Transactions on 13*, 6 (2007), 1262–1269.
33. Tory, M., Swindells, C., and Dreezer, R. Comparing dot and landscape spatializations for visual memory differences. *Visualization and Computer Graphics, IEEE Transactions on 15*, 6 (2009), 1033–1040.
34. Tufte, E. R., and Graves-Morris, P. *The visual display of quantitative information*, vol. 2. Graphics press Cheshire, CT, 1983.
35. Ware, C. *Information visualization: perception for design*. Morgan Kaufmann, 2012.