Jianu, R., Rusu, A., Hu, Y. & Taggart, D. (2014). How to Display Group Information on Node-Link Diagrams: An Evaluation. IEEE Transactions on Visualization and Computer Graphics, 20(11), pp. 1530-1541. doi: 10.1109/TVCG.2014.2315995

How to Display Group Information on Node-Link Diagrams: an Evaluation

Radu Jianu, Adrian Rusu, Yifan Hu, and Douglas Taggart

Abstract— We present the results of evaluating four techniques for displaying group or cluster information overlaid on node-link diagrams: node coloring, GMap, BubbleSets, and LineSets. The contributions of the paper are three fold. First, we present quantitative results and statistical analyses of data from an online study in which approximately 800 subjects performed ten types of group and network tasks in the four evaluated visualizations. Specifically, we show that BubbleSets is the best alternative for tasks involving group membership assessment; that visually encoding group information over basic node-link diagrams incurs an accuracy penalty of about 25% in solving network tasks; and that GMap's use of prominent group labels improves memorability. We also show that GMap's visual metaphor can be slightly altered to outperform BubbleSets in group membership assessment. Second, we discuss visual characteristics that can explain the observed quantitative differences in the four visualizations and suggest design recommendations. This discussion is supported by a small scale eye-tracking study and previous results from the visualization literature. Third, we present an easily extensible user study methodology.

Index Terms—networks, sets, clustering, evaluation, user study.

I. INTRODUCTION

VISUALIZATION of connectivity and relational data as node-link diagrams has demonstrated its effectiveness as an analysis tool in a wide range of application domains such as social sciences, intelligence analysis, engineering, computer science, and diverse biology related research areas. However, as noted by Collins et al. [7], connectivity is rarely the only type of information used in these domains. Instead, connectivity information is often meshed with set, group, or cluster information, order and quantitative information, and domain specific metadata. To exemplify, protein interaction visualizations often benefit from highlighting groups of proteins that are co-activated or that share similar function [29,4].

Here we present a systematic evaluation of four techniques for meshing connectivity data with group information. Specifically, in an online Mechanical Turk study, we evaluated four visualizations for overlaying group information on node-link diagrams and used the resulting data to suggest design guidelines. The four visualizations were the GMap algorithm proposed by Gansner et al. [14,12], the BubbleSets algorithm introduced by Collins et al. [7], the LineSets approach of Alper et al. [1], and, as a baseline, a standard node-link diagram using colored labels to encode group data. We chose these four because they represent recent state of the art techniques that are capable of capturing groups that are not necessarily spatially co-located, a property which renders them highly versatile. We note that our study only evaluates groups that are not overlapping. We use the term group or cluster to emphasize the distinction from the more general concept of sets.

Our study found that there are meaningful differences in how people perform tasks in these four visualizations. These differences are linked to visual encodings specific to each method. For instance, unlike node coloring and LineSets, GMap and BubbleSets are background filling, thereby providing more colored surface that can be used for group membership assessment. The contiguous contours employed by BubbleSets and the connecting line in LineSets help disambiguate whether similarly colored nodes or areas are part of the same group, but require additional cognitive and perceptual resources in doing so. Finally, overlaying a visual layer to encode group information seemed to reduce people's ability to perform network related tasks such as path tracing by about 25%.

The significance of our work is two-fold. First, mapping visualizations to tasks that they address well can lead to more effective visualization deployment. Second, understanding how visual encodings support or inhibit data reading can inform future design. To the best of our knowledge there are few previous results on evaluating techniques for displaying group information that is not spatially co-located, let alone in conjunction with node-link diagrams. Moreover, we augment the body of general evaluative visualization research.

Contributions: (i) a quantification of the effectiveness of four techniques for displaying group information overlaid on node-link diagrams for ten types of tasks; (ii) a discussion of visual attributes that may explain the observed differences in the four visualizations; (iii) a set of design recommendations; (iv) an easily extensible user study methodology.

II. RELATED WORK

Several evaluation results on network visualization exist. Specifically, Purchase et al. [35, 36, 37] looked at how graph drawing aesthetics like edge crossings and symmetries impact

- R.J. Author is with The School Of Computing and Information Sciences, Florida International University, 11200 SW 8th Street, ECS354, Miami, FL 33199. E-mail: rdjianu@ cis.fiu.edu.
- A.R. Author Jr. is with the Department of Computer Science, Rowan University, 201 Mullica Hill Road, Glassboro, NJ 08028. E-mail: rusu@rowan.edu.
- Y.H Author is with the AT&T Labs Research, 180 Park Ave., Florham Park, NJ 07932. E-mail: yifanhu@research.att.com.
- D.T Author is with the Department of Computer Science, Rowan University, 201 Mullica Hill Road, Glassboro, NJ 08208. E-mail:dougtag@gmai.com.

people's abilities to perform graph reading tasks such as path tracing. Ware et al. have done similar work [43], extended these studies to 3D [45, 46], and proposed low level perceptual models of edge tracing [44]. Huang et al. have used eye-tracking and user studies to understand visual network perception [21,20,22,23]. Archambault, Ghani, and Farrugia have evaluated perceptual characteristics and memorability in dynamic, animated graphs [15, 11, 3, 2]. Finally, Marriott et al. [33] investigate the memorability of graph features as a proxy for understanding how graphs are mentally represented. While these are valuable results and advance our understanding of how visual encodings contribute to visualization effectiveness, evaluative work in graph visualization and visualization is still lagging behind technique development. To exemplify, the latest graph drawing survey [42], cites approximately 100 technique and algorithm papers but only about 30 design and evaluation studies combined. The work presented in this paper helps bridge this gap.

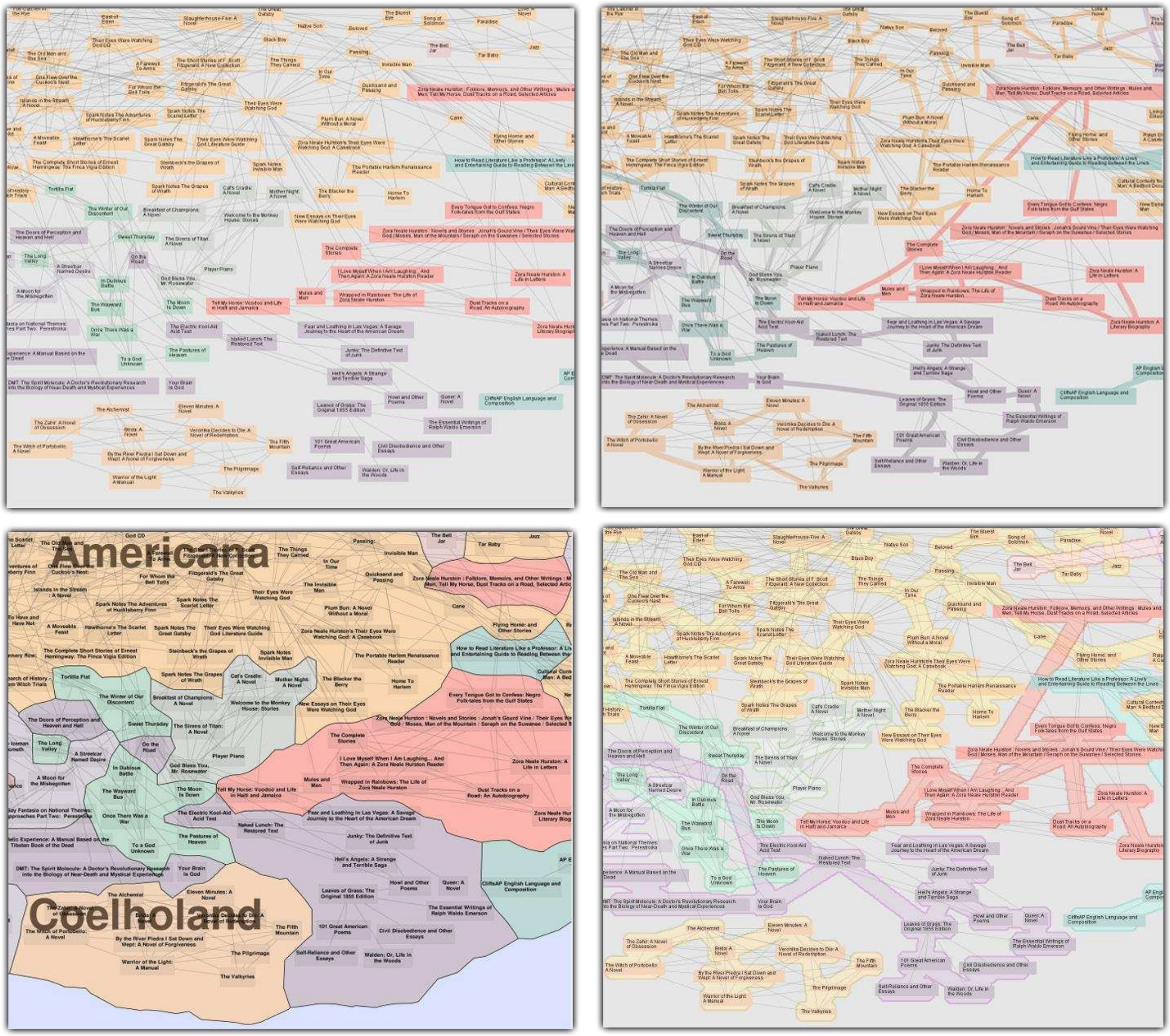


Fig. 1: Four visualizations for viewing group information over node-link diagrams: colored node-link diagram (top left), LineSets (top right), GMap (bottom left), and BubbleSets (bottom right).

We evaluate four visualization methods for overlaying group information over node-link diagrams. First, GMap, introduced by Gansner et al. [12,14], uses a space filling map metaphor to enclose group members into "countries", "seas", and "lakes". Second, BubbleSets, introduced by Collins et al. [7] and inspired by Heine and Scheuermann's "blob" approach [19], draws contiguous contours around nodes of the same group even if the nodes are not spatially co-located. Third, LineSets, introduced by Alper et al. [1], links all members of a set with a continuous curve of distinct color. BubbleSets and

LineSets are set visualizations, meaning they can display overlapping groups. Finally, as a base method, we use simple node coloring of a standard node-link diagram to make groups apparent. These four visualizations are demonstrated in Fig. 1 and detailed in section 3.

Other techniques for visualizing group or set information exist, but are less apt at handling scattered and intertwined group members with predefined layouts [7]. Examples include the untangled Euler diagrams [38], the visualization of overlapping sets by Simonetto et al. [41], and other methods involving the drawing of convex hulls around group members [18,10, 19].

In terms of methodology and contributions our work comes close to recent work by Kong et al on tree-maps [31]. Like them, we perform an online comparative study and derive design principles and guidelines from our observations and data analyses. Our work also comes close to evaluative work by Rogers et al. on visual properties of Euler diagrams [39,6,5]. Similarly to them, we find that subtle variations in properties of group visualizations can lead to differences in task performance. The online evaluation approach for visualization, specifically using the Mechanical Turk service, has been described by Heer and Bostock [17], and has demonstrated its benefits in a range of recent visualization work [34, 16]. Finally, we note that in our selection of network tasks that we evaluate, we used the task taxonomy described by Lee et al. [32].

III. METHODS

The quantitative results presented here are distilled from data collected in an online study. We ran approximately 800 unique subjects in a between-users design, tested ten different tasks categories with multiple task instances per category, and evaluated four visualizations. Subjects solved three or four task categories in sessions of around five to ten minutes and could not participate in more than one session. The study was distributed online using the Mechanical Turk service. Users were presented with a brief introduction to their assigned visualization, were given one minute to explore it, and were then guided through a series of short tasks ranging between 2 seconds and 45 second, organized into three or four task-categories. Results were collected automatically and stored in a database.

The following sections provide details about the design of the study. Specifically, we give details about our participants, describe the four evaluated visualizations, the tasks, the study design, and the online distribution. Details about the numbers of users for each task, numbers of task instances in each category, and task duration are summarized in Table 1.

A. Participants

Participants were recruited on the Amazon Mechanical Turk website. As such, participant demographics correspond to that of this online service [40].

We gathered data from a total of 788 unique users. The way they were distributed over tasks and conditions is summarized in Table 1 and section 3.2. The participant numbers listed in Table 1 are approximate and were the ones we commissioned on Mechanical Turk. In reality we gathered data from slightly more participants (2-5 /condition). This happened because some users did not complete the task in the specified amount of time, therefore could not officially submit their result and be counted by Mechanical Turk. However, their full results were stored in our database and they were compensated appropriately. A second reason why participant numbers are listed as approximate is because we discarded results from 12 users who selected identical answers for all tasks in a task category. No other heuristics were used to clean the data.

B. Evaluated Visualizations

The four evaluated techniques were Gasner et al.'s GMap algorithm [14, 12], Collins et al.'s BubbleSets [7], the LineSets approach of Alper et al. [1], and, as a base condition, simple node coloring. The four visualizations are exemplified in Fig. 1.

GMap uses a map metaphor and partitions the canvas into "countries", "seas", and "lakes" that correspond to groups of nodes. Interestingly, two visually disjointed areas can still be part of the same group. The groups border each other directly without leaving blank space in between. Moreover, GMap draws "country" names over larger or more distinctive groups. This algorithm cannot handle overlapping groups (i.e. is not a set visualization).

In BubbleSets, enclosing contours (bubbles) are drawn around groups of nodes and the enclosed area is colored distinctively. BubbleSets handles groups that are spatially dispersed by computing an invisible skeleton of edges that connect the disjointed group areas, and by then drawing bubbles around that skeleton. The tightness of the bubble around elements and connective edges can be controlled through a parameter: low values will result in "skinny" bubbles such as those seen in Fig. 1; high values will produce visual results that are closer to GMap: blown up bubbles that fill the space without much overlap. We chose to evaluate "skinny" bubbles to exploit the more significant difference from GMap and thus provide a better coverage of the visual design space. Unlike GMap, BubbleSets is a set visualization and can handle overlapping groups.

In LineSets, nodes are colored distinctively to indicate membership and a curved band sharing the group's color passes through all members and links them visually. The order in which elements are visited by the curve determines the shape of the curve and may influence the effectiveness of the visualization. Relying on data from user experimentation, the authors of LineSets settle on sequences produced by the traveling salesman algorithm. We used this same heuristic in our evaluated implementation. LineSets is also a set visualization.

Finally, as a base case, we used a traditional node-link diagram and employed node coloring to display group information. It is impossible to represent overlapping groups using this technique.

These four visualizations share pairwise similarities but also cover a significant portion of visual design space. Viewed globally, they span a continuum over several visual coordinates. LineSets changes the colored nodes representation only incrementally. It adds the property of group contiguity and increases the amount of color assigned to groups. BubbleSets maintains group contiguity but represents it differently. Its connecting skeleton facilitates shorter connections between group members when compared to the non-branching curve used in lines. Furthermore, BubbleSets increases the amount of color assigned to groups. In some cases nearby group members become grouped together into unitary blobs, thereby introducing the notion of a colored region. Finally, GMap maximizes the size and local contiguity of colored group areas by using a completely filling approach. At the same time however, it abandons the idea contiguity.

During study design we considered the option of removing group labels from the GMap visualization. Ultimately, we decided against it as they seem to be an intentional and integral part of the GMap technique and not of the other visualizations. At the end of section 4.3 we present two post-hoc analyses that look into possible impacts of labels on the evaluated tasks.

The underlying data used for the evaluated visualizations was a book-recommendation network of approximately 850 books. Two books were linked together if customers had purchased both of the two books. The book groups used in the study and displayed in the figures were created by running a graph-clustering algorithm. Essentially, book groups corresponded to books often purchased together. The dataset contained at total of 26 sets. As mentioned before, groups were non-overlapping – a book belonged to just one group. All three visualizations used the same underlying network layout which was produced using the neato algorithm [13].

All visualizations used fully visible labels fitted to book title lengths. Label sizes were identical between the three evaluated visualizations. We also controlled the coloring scheme used to distinguish between groups by using identical coloring in all three conditions. It is worth noting that while GMap, colored nodes, and LineSets use opaque coloring, BubbleSets uses transparency. We factored this into our design and assigned bubble-colors that cancelled out the transparency to give identical colors to the other conditions when viewed over blank background. Finally, the visualizations used the same shade of grey to draw network edges.

C. Evaluated Tasks

We evaluated ten different types of tasks falling broadly in three categories: group tasks, network tasks, and mixed group-network tasks. Each task was evaluated with multiple instances. We note that network tasks were inspired by the graph task taxonomy described in [32]. We have not considered tasks involving overlapping groups, a limitation discussed in section 5.

Task descriptions and details are provided in the left column of Table 1. The first two tasks tested the ability of users to tell whether two nodes belong to the same group or not. Examples of such tasks are shown in Fig 2. The next two tasks test the ability of users to perceive the overall structure and properties of groups. The following three tasks test subjects' ability to solve network specific tasks (Fig 3). Another two tasks tested subjects' ability to perceive and consider group and network properties at the same time. Finally, the last task was designed to test how the visualizations impact the ability of users to remember node positions. This last task was tested in two stages. First, at the beginning of a study session, subjects were asked to search for three book titles using the search feature. Once they found the books, subjects had to retrieve and record several pieces of information about them. At the end of the study, after approximately five minutes, subjects were asked to relocate the same three books without the assistance of the search tool.

The number of task instances per task type is specified in Table 1. We tried to run as many instances as possible to capture various visual scenarios and increase the chance of capturing significant differences. At the same time we balanced that with the requirement to keep the study under ten minutes.

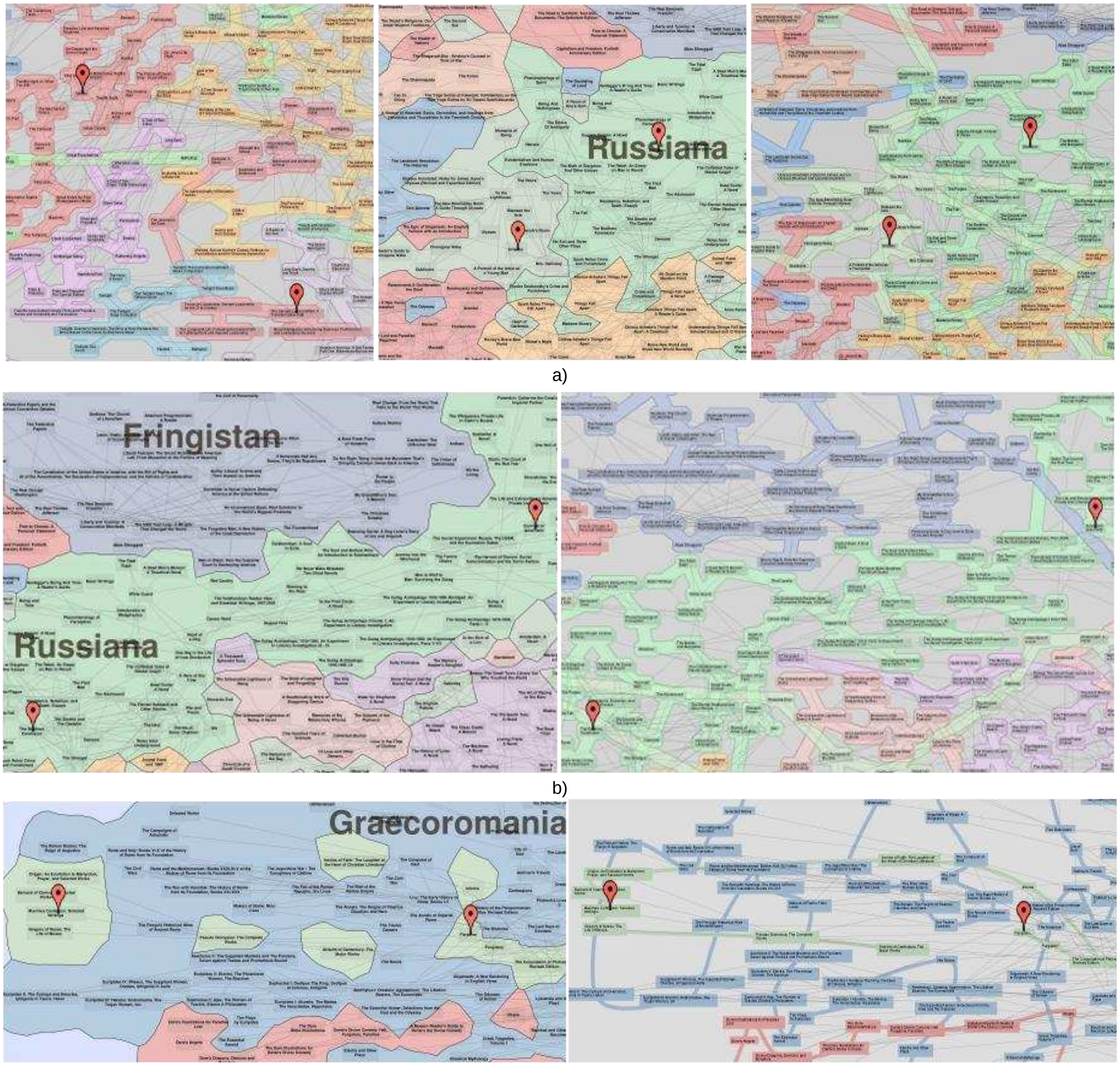
All tasks were timed at durations listed in Table 1. These times were established by first measuring how fast the authors of the study could complete the tasks. In a second stage three graduate students unaffiliated with the project were piloted to test the study design before deployment. Times were further adjusted based on their performance.

Times were chosen deliberately to be short and challenging rather than represent a cutoff point that users rarely reach. This is similar to enforced time limits used in perceptual studies to ensure that participants solve tasks using perception and estimation rather than deliberation. In our opinion, the power of visualization comes from allowing users to perform data reading tasks fast and at a global scale, using their perception, intuition, and fast eye movements. Accuracy is important if it can be achieved faster than running explicit data queries to solve the tasks. To exemplify, if our user require more than seven seconds determining whether two points belong to the same group, they would probably be better off selecting the points and asking the computer to provide the correct answer.

D. Online study

Our main study was online and between users. We evaluated the ten types of tasks in four separate experimental sessions: tasks 1 and 3 were evaluated in a first session; tasks 2 and 4 in a second; tasks 5, 6 and 7 in a third; and the last three tasks in a fourth. To control for learning effects, subjects were allowed to participate in the study only once. As such, each subject performed only one of these task combinations using only one of the four visualizations.

These design choices were motivated by the desire to keep experimental sessions short and amenable to online distribution using the Mecahnical Turk service [17]. Our experimental sessions ended up lasting between five and ten minutes.

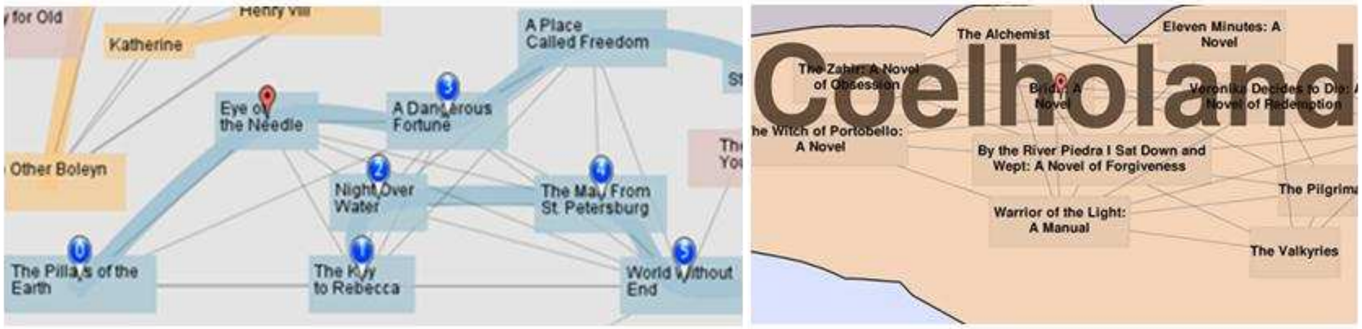


a)

b)

c)

Fig. 2: Example of group membership tasks.



a) b)

Fig. 3: Example of network tasks

Each experiment started with a brief introduction to the purpose of the study, to the data, and to the visualizations used. We used non-technical language such as books instead of nodes and connections instead of edges throughout the study and explained concepts such as groups or paths whenever necessary.

Following the introductory briefing, subjects were allowed to experiment with the visualization for one minute. They were encouraged to search for and select nodes, zoom and pan, explore node attributes, and understand the structure of the visualization.

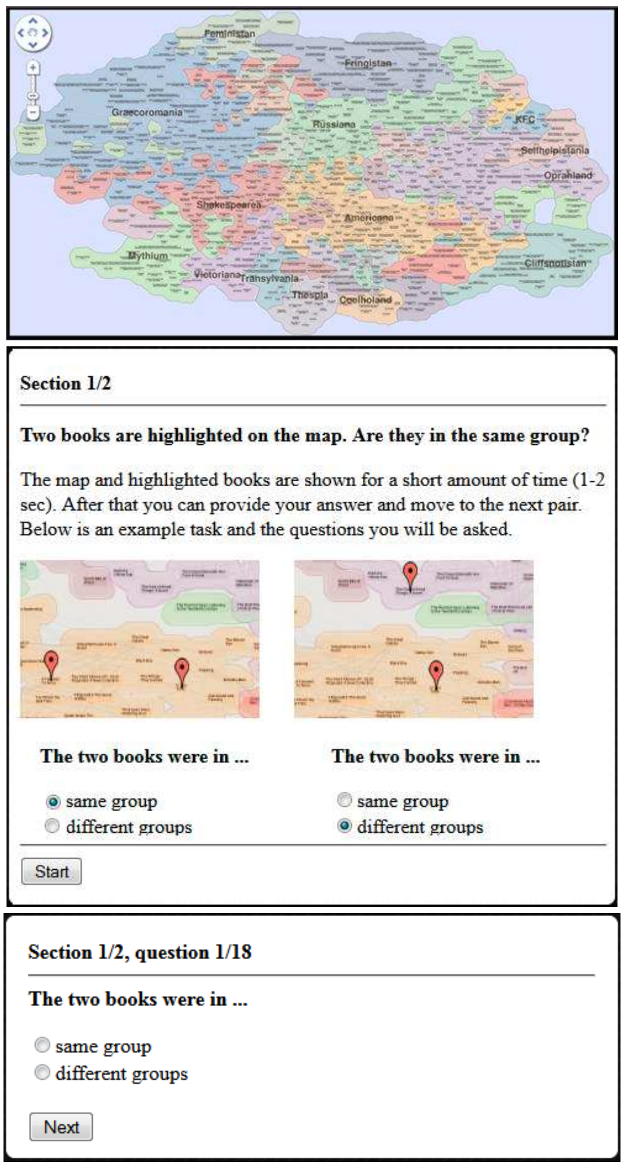


Fig. 4: Timed browser dialogues guide the user through the online study by providing task instructions and collecting answers.

Subjects were then introduced to their assigned tasks. Each task type was preceded by a short instructional page complete with visual examples of both the task and the possible responses (Fig. 4). Subjects were then guided through multiple instances of the task, each followed by a modal window which recorded their responses (Fig. 4). Each task was timed and the count-down timer was visible to subjects. At the beginning of each task instance, the map was automatically translated and zoomed to provide the best possible view of the task and thus minimize the cost of interaction.

The study was run online using the Mechanical Turk service following the approach demonstrated by Heer and Bostock [17]. To this purpose the four network visualizations were rendered as large static images, tiled, and distributed as Google Maps (Fig. 4). Using Google maps to distribute non-geographic data visualization has been used before in [27, 26]. This provides a familiar and intuitive exploration framework for users and ensures quick loading and interaction times. The Google maps were augmented with search functionality and node selection capabilities. The jquery and colorbox toolkits were used to create autocomplete search functionality and timed modal windows that guided users through the study (Fig. 4).

In retrospect, one drawback of our study was the failure to include control tasks to identify which users are solving the tasks and which are simply selecting random answers (also see section 3.1). We believe control tasks would have contributed to cleaner results. Even so, randomness would have necessarily affected all conditions equally and is as such captured in our statistical analysis. As shown by the boxplots in Table 1, subjects exhibited fairly consistent behavior between and within conditions for most tasks, with one or two exceptions. Moreover, the similarity between results obtained in the first two experiments, with different subjects, supports the validity of the results.

E. Eye-tracking study

We followed our online study with an informal, small scale eye-tracking study aimed at helping us explain some of the effects seen in the quantitative results. Eye-tracking can be helpful in this endeavor since gaze patterns are correlated with task intention [9].

Specifically, four subjects were asked to solve tasks 1 and 2 using the same conditions as in the online study. The equipment used was a RED 120Hz eye tracker from Sensory Motor Instruments. The collected data was visualized as gaze heatmaps and scan-paths and analyzed qualitatively. This methodology is similarly to work by Huang et al. [20]. However, we note that our four subjects are significantly fewer than Huang's thirteen subjects. Our eye-tracking investigation was meant to inform the discussions in section 4 rather than constitute a user study in its own right.

IV. Results

We first summarize the quantitative results of our study and then discuss them. These results are fully listed in Table 1. We then summarize visual attributes specific to each of the three evaluated techniques. We do so by using insights from our eye-tracking study, and previous findings from the visualization literature.

A. Quantitative results

The quantitative results of our comparative study are captured in Table 1. Since tasks were timed and constant between conditions, we used accuracy as a sole measure of performance. However, since we chose times that are borderline short for each task, decreased accuracy could be attributable to more time being required. In this light, our accuracy measure is a combined measure.

For each type of task, an error specific to that task was computed for each task instance, and then averaged over all task instances for each user. The aggregated results of all users, averages with standard error bars and boxplots, are depicted in the middle and right columns of Table 1. For all ten tasks, the computation and then aggregation of errors lead to discrete rather than continuous results. For instance, for tasks 1 and 2, the reported responses where either yes or no, the errors per task instance either 0 or 1, and the average errors per user between 0 and 1 with increments that depended on the number of task instances (1/18 and 1/12). Also, the boxplots indicate that for most tasks the result distributions for the four visualizations are similar even if not identical. As such the statistical method chosen for analysis was Kruskal Wallis. A Mann-Whitney posthoc analysis with Bonferroni correction was then used to distil out differences between the four visualizations. All analyses were done using R.

B. Result discussion

Group tasks: Overall, the results on all four group tasks suggest that BubbleSets users were most accurate, with LineSets users in close second. Specifically, BubbleSets outperformed colored nodes in three of the four task categories, outperformed GMap in two of the four categories, and outperformed LineSets in just one of the four task categories, while not itself being outperformed in any of the tasks. At the same time, colored nodes appears to be the least effective method as it is outperformed by at least one other technique in each of the four tasks.

Two factors support the validity of the results. First, the results are fairly consistent over the four tasks, even though pairs of tasks were run separately with different user groups (tasks 1 and 3, and tasks 2 and 4). Second, the relative results between BubbleSets and LineSets are similar to those reported by LineSets authors.

It's worth noting that the difference between BubbleSets and LineSets in tasks 3 and 4 are not significant. Both these tasks tested users' ability to perceive groups as a whole rather than determine set membership in keyhole scenarios. Combined with the fact that performance was also the same in task 1, LineSets is likely to be a viable alternative to BubbleSets in real-life scenarios where users are exposed to a visualization long periods of time.

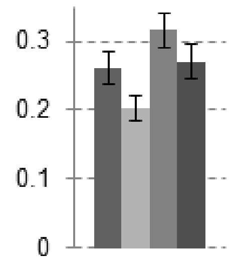
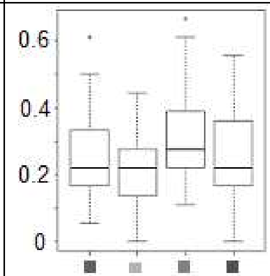
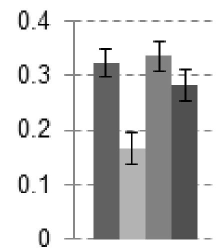
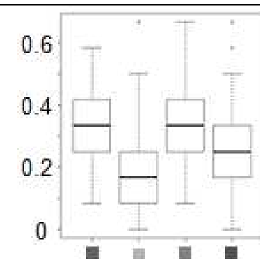
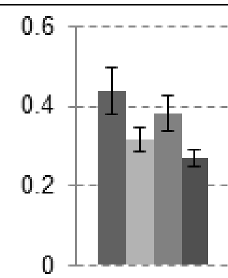
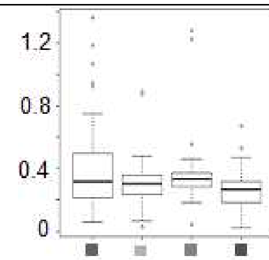
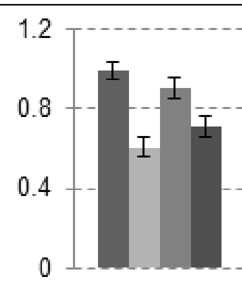
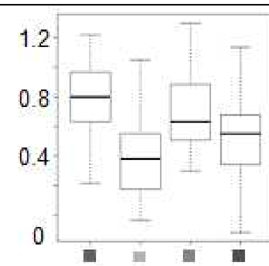
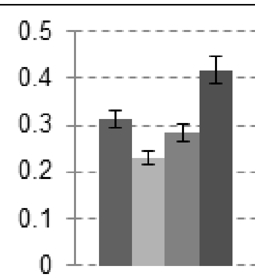
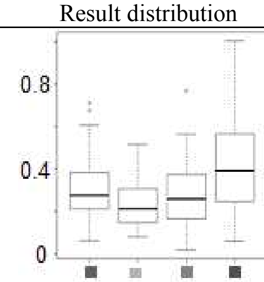
A more detailed look into the results of task 1, combined with an analysis of eye-tracking data, leads to further insights into the visual attributes that support or inhibit membership assessment. The eighteen instances of task 1 could be roughly divided in four categories. In a first category the two highlighted nodes were in well separated groups with sometimes similar color (Fig. 2a); in a second category they were in a continuous set (Fig. 2b); in a third category they were in adjacent groups, sometimes with similar color (Fig. 2a, middle and right); in the fourth category they were in spatially scattered node clusters that belonged to the same group (Fig. 2c). Figure 5 shows task 1 results split over these four categories. GMap outperforms all other techniques in two of the four task categories (categories one and three, $p<0.001$), but performs poorly in the fourth category ($p<0.001$). No difference between techniques could be found for category two tasks. Eye-tracking provides further evidence for GMap superiority in certain scenarios: users occasionally seemed to have made up their mind even before the short 2 seconds expired, as shown by an unusually long last fixation on no particular visual element.

It is easy to explain the negative GMap results in the fourth category tasks. Since there is no explicit connection between disconnected "islands" of the same set users are unsure whether they belong to the same group or not. This was also stated by our eye-tracking participants and is likely the reason why GMap cannot compete with BubbleSets and LineSets in accuracy.

We hypothesize that the positive results of GMap can be attributed to the large, continuous color areas assigned to groups. While lines or thin connective bubbles can help disambiguate group structures, they require precise and deliberate attention. Instead, GMap users can use their peripheral vision to trace across large, identically colored areas. Moreover, the additional area attributed to groups in GMap makes it easier to accurately compare colors [30]. This may also explain why users performed poorly in the colored nodes condition, which offers the least amount of color area.

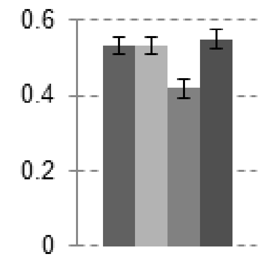
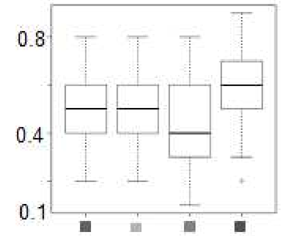
Eye-tracking data supports these hypotheses. First, GMap users seemed to fixate in the general vicinity of highlighted nodes, but not necessarily on the node, indicating that users could rely on their peripheral vision to match a node to a group. Furthermore, fixations in GMap seemed generally shorter and users rarely revisited previously fixated points. At the opposite end, fixations in the colored nodes condition were generally right on top of highlighted nodes and revisited previously fixated points often. Characteristic of fixations in both the BubbleSets and LineSets conditions were that they often dwelled on visually complex areas such as overlapping or branching bubble areas or intersecting lines. Moreover, initial highlighted node fixations in BubbleSets were similarly short to GMap, suggesting that color area helped in this case as well.

TABLE 1

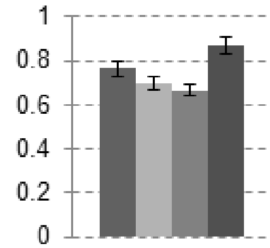
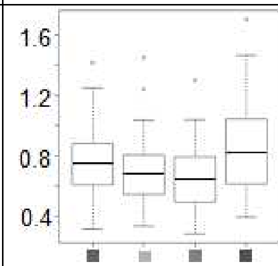
<table>
<tr><td colspan="3">gmap bubble node lines</td></tr>
<tr><td colspan="3">Group Tasks</td></tr>
<tr><td>Task and measure</td><td>Result means</td><td>Result distribution</td></tr>
<tr><td><u>Task 1, Intuitive group membership:</u> given two highlighted nodes and little time, subjects judge if the nodes are in the same group.
task instances: 18; # users: ~30 / visualization; task duration: 2 sec.
Measure: Errors (0 – correct answer, 1 – incorrect answer) were averaged over 18 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=10.58$, $p < 0.05$
Posthoc (Mann-Whitney with Bonferroni correction): BubbleSets vs. colored nodes ($p<0.01$, $r=0.40$)</td><td></td><td></td></tr>
<tr><td><u>Task 2, Deliberative group membership:</u> given two highlighted nodes and sufficient time (7sec), subjects judge if the nodes are in the same group.
task instances: 12; # users: ~30 / visualization; task duration: 7 sec.
Measure: Errors (0 – correct answer, 1 – incorrect answer) were averaged over 12 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=23.37$, $p < 0.001$
Posthoc (Mann-Whitney with Bonferroni correction): GMap vs. BubbleSets ($p<0.001$, $r=0.5$); BubbleSets vs. colored nodes ($p<0.001$, $r=0.54$); BubbleSets vs. LineSets ($p<0.05$, $r=0.36$)</td><td></td><td></td></tr>
<tr><td><u>Task 3, Number of sets:</u> given a visible graph region, subjects counted the number of distinct groups in the region.
task instances: 5; # users: 31 / visualization; task duration: 20 sec.
Measure: Errors (|true_#sets-reported_#sets| / true_#sets) were averaged over 5 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=8.62$, $p < 0.05$
Posthoc (Mann-Whitney with Bonferroni correction): colored nodes vs. LineSets ($p<0.001$, $r=0.34$)</td><td></td><td></td></tr>
<tr><td><u>Task 4, Relative group size:</u> given two highlighted groups, subjects judge their relative size difference (1X, 1.5X, 2X, 3X bigger).
task instances: 7; # users: ~30 / visualization; task duration: 20 sec.
Measure: Errors (|true_size_ratio-reported_size_ratio|) were averaged over 5 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=28.67$, $p < 0.001$
Posthoc (Mann-Whitney with Bonferroni correction): GMap vs. BubbleSets ($p<0.001$, $r=0.58$); GMap vs. LineSets ($p<0.01$, $r=0.45$); BubbleSets vs. colored nodes ($p<0.01$, $r=0.45$)</td><td></td><td></td></tr>
<tr><td colspan="3">Network Tasks</td></tr>
<tr><td>Task and measure</td><td>Result means</td><td>Result distribution</td></tr>
<tr><td><u>Task 5, Node degree estimation:</u> given a highlighted node, subjects determine its degree.
task instances: 7; # users: ~60 / visualization; task duration: 7-13 sec., depending on degree
Measure: Errors (|true_node_deg-reported_node_deg|/true_node_deg) were averaged over 7 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=30.07$, $p < 0.05$
Posthoc (Mann-Whitney with Bonferroni correction): GMap vs. BubbleSets ($p<0.05$, $r=0.27$); GMap vs. LineSets ($p< 0.05$, $r=0.24$); BubbleSets vs. LineSets ($p<0.001$, $r=0.45$); colored nodes vs LineSets ($p<0.01$, $r=0.30$)</td><td></td><td></td></tr>
</table>

<table>
<tr><td>

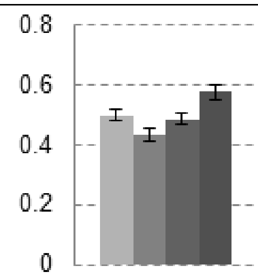
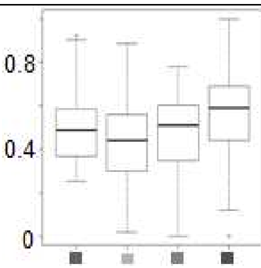
<u>Task 6, Path tracing:</u> given a sequence of nodes, subjects determine if the sequence is a valid path (edges between consecutive nodes are present).
task instances: 10; # users: ~60 / visualization; task duration: 8-17 sec., depending on task difficulty
Measure: Errors (0 – correct answer, 1 – incorrect answer) were averaged over 10 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=15.18$, $p < 0.01$
Posthoc (Mann-Whitney with Bonferroni correction): colored nodes vs. GMap ($p<0.05$, r=0.30); colred nodes vs. BubbleSets ($p<0.05$, r=0.30); colored nodes vs. LineSets ($p<0.01$, r=0.35)

</td><td>
</td><td>
</td></tr>
<tr><td>

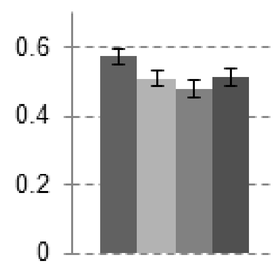
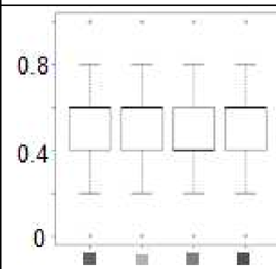
<u>Task 7, Neighbors selection:</u> given a highlighted node, subjects select all its neighbors.
task instances: 3; # users: ~60 / visualization; task duration: 15-45sec., depending on task difficulty
Measure: Errors ((false_positives+false_negatives)/true_#neighbors) were averaged over 3 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=18.10$, $p < 0.001$
Posthoc (Mann-Whitney with Bonferroni correction): BubbleSets vs. LinesSets ($p<0.01$, r=0.28); colored nodes vs. LineSets ($p<0.001$, r=0.34)

</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Combined Group-Network Tasks</td></tr>
<tr><td>Task and measure</td><td>Result means</td><td>Result distribution</td></tr>
<tr><td>

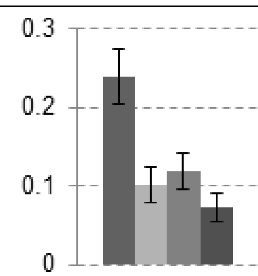
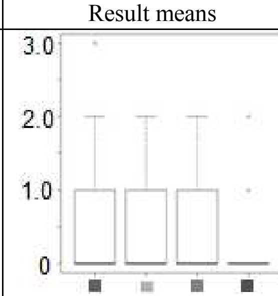
<u>Task 8, Highest degree node:</u> given a highlighted group, subjects identify the highest degree node in that group.
task instances: 4; # users: ~70 / visualization; task duration: 30 sec.
Measure: Errors (1 - reported_highest_deg/true_heighest_deg) were averaged over 4 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=19.86$, $p < 0.001$
Posthoc (Mann-Whitney with Bonferroni correction): LineSets vs. GMap ($p < 0.05$, r = 0.22); LineSets vs. BubbleSets ($p < 0.001$, r = 0.34); LineSets vs. colored nodes ($p<0.05$, r=0.23)

</td><td>
</td><td>
</td></tr>
<tr><td>

<u>Task 9, Tracing paths over groups:</u> given a sequence of highlighted nodes, subjects determine if the sequence is a valid path (edges between consecutive nodes are present), and if no two consecutive nodes are in the same group.
task instances: 5; # users: ~70 / visualization; task duration: 20-25 sec., depending on task difficulty
Measure: Errors (0 – correct answer, 1 – incorrect answer) were averaged over 5 task instances for each user and then averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=8.11$, $p < 0.05$
Posthoc (Mann-Whitney with Bonferroni correction): GMap vs. colored nodes ($p < 0.05$, r = 0.22).

</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Memory Task</td></tr>
<tr><td>Task and measure</td><td>Result distribution</td><td>Result means</td></tr>
<tr><td>

<u>Task 10, Memory:</u> given a node that is the object of several tasks, subjects attempt to locate the node several minutes later without the assistance of search tools.
task instances: 3; # users: ~70 / visualization;
Measure: Instance of recollected nodes in the three tasks for each were averaged over all users for each of the four visualizations.
Significance: Kruskal Wallis: $\chi^2(3)=16.87$, $p < 0.001$
Posthoc (Mann-Whitney with Bonferroni correction): GMap vs. BubbleSets ($p < 0.05$, r = 0.23); GMap vs. LineSets ($p < 0.01$, r=0.31)

</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">■ gmap ■ bubble ■ node ■ lines</td></tr>
</table>

Network tasks: The result of task 6, path tracing, confirmed our hypothesis that the node condition would perform best. Moreover, the result quantifies the accuracy penalty incurred by adding visual elements that encode group information to about 25%. Interestingly there were few differences between GMap, BubbleSets, and LineSets.

We also believe the results in tasks 5 and 7 illustrate two effects. First, degree estimation in GMap was hampered in two cases by group labels (Fig. 3a), explaining why GMap users performed worse on task 5 than BubbleSets and colored nodes users. Second, LineSets users seemed to judge that two nodes are connected if they are visited by the same set curve. This was noticeable since many users in the line condition selected as neighbors those nodes that were visited by the set curve shortly before or after the highlighted node. This misunderstanding explains the low performance of line users in tasks 5 and 7. Interestingly however, this did not seem to impact performance in the path tracking task even though this task was performed between tasks 5 and 7. We hypothesize that if these two effects were removed, performance in tasks 5 and 7 would be similar across visualizations.

Mixed group-network tasks: Results in tasks 8 and 9 seem to combine effects observed in group tasks and network task. In task 8, finding the highest degree node, LineSets performs worst, similarly to results seen in the degree estimation task. In task 9 the trend of the means is similar to that in task 6, with the node condition showing slightly better accuracy means. However, differences between means are small and a significant difference was only found between GMap and colored nodes. We hypothesize that the insignificant differences between means occur because the better performance of BubbleSets and LineSets in group tasks offsets the advantage of colored nodes in path tracings tasks.

Memorability: GMap authors claim that their visualization leverages people's intuition of working with maps [14, 12]. Our data provides support for this claim by showing that subjects were indeed more likely to remember node location in GMap than in the other three visualizations. However, as shown at the end of section 4.3, the improved memorability is attributable to the addition of group labels rather than the visual layout of GMap.

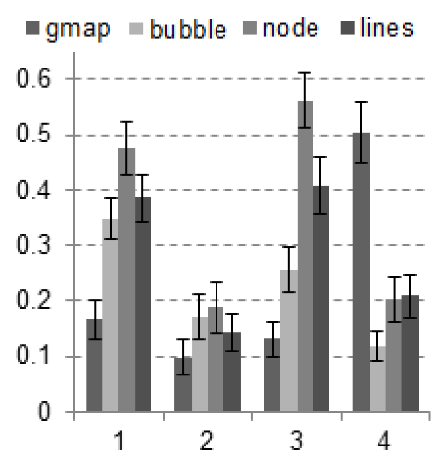
C. Visual attributes

Here we summarize and discuss a few visual attributes of the four visualizations. The discussion can inform future designs and lead to tools that are better suited for the tasks they aim to support.

Connections between dispersed group members are used in both BubbleSets and LineSets. As discussed, they help define the structure of a group but require deliberate and focused tracing, which is time consuming. We hypothesized that the method employed by BubbleSets is superior to that of LineSets because it favors shorter visual connections between group members. Since the curve in LineSets visits nodes sequentially, tracing it between nodes would be laborious. However, we found no eye-tracking evidence that LineSets users actually trace the curve between highlighted nodes. Instead, they identify areas where the curve wraps around many group members, and identify them holistically as part of a group. They then try to find a curve segment that unites two islands containing highlighted nodes. This process is similar to what happens in BubbleSets and probably contributes to close results observed between the two techniques in group tasks.

Color area: The ability to distinguish between colors depends on the amount of colored area at a users' disposal [30]. We believe that GMap has an advantage in this respect. A qualitative inspection of our results also suggested that performance in the colored nodes condition was dependent on label sizes and highly susceptible to errors whenever colors where similar. BubbleSets and LineSets are a mid-solution, with BubbleSets providing more color. In fact, we believe that an important advantage of LineSets is in increasing the amount of color displayed, in addition to connecting same-set elements.

Color continuity: For colored nodes, users have to perform a discrete comparison between small color patches by moving their gaze back and forth between the nodes. Conversely, in GMap and BubbleSets, and to some degree in LineSets, users can often trace the continuous area from one node to the other to determine group membership. As discussed, evidence for this was found in our eye-tracking data.



1: $\chi^2(3)=25.27$, $p<0.001$ (GMap vs All, $p<0.001$)
2: $\chi^2(3)=3.5$, $p>0.05$
3: $\chi^2(3)=37.79$, $p<0.001$ (GMap vs. All, $p<0.001$)
4: $\chi^2(3)=33.17$, $p<0.001$ (GMap vs All, $p<0.001$)

Figure 5: Task 1 results split in four scenario categories

Visual clutter: The four methods introduce various types of visual clutter which affect data reading. GMap introduces clutter by displaying labels over distinctive sets. While perhaps useful in separating groups and aiding memorization, these labels can occlude visualization elements, as discussed in section 4.2. BubbleSets introduces clutter around areas where multiple groups overlap. LineSets introduces clutter due to the winding curves. Finally, it may appear that the node condition features the least visual clutter. However, in colored nodes, discrete entities have to be visually aggregated into groups, a task readily done by the other visualizations, particularly GMap. This drawback was indirectly observed in the data. Group membership errors in the colored nodes condition were small when groups were large and cohesive, thereby facilitating the perception of the nodes as a single unit. Conversely, when nodes from many clusters where interspersed, colored nodes showed very poor performance.

Labels: GMap employs explicit group labels while all other visualizations do not. To investigate the degree to which these labels may have interfered with tasks one through nine, we searched for all task instances in which labels directly overlap with nodes involved in the task. Such cases represent approximately 15% of all task instances. We then removed these cases and re-plotted the bar charts in table 1. Charts remained virtually unchanged except for the degree estimation chart, a case already discussed in section 4.2. This informal post-hoc analysis does not dismiss the hypothesis that labels had a major impact on tasks one through nine but supports the case against it.

To investigate the effect of labels on the memorability task, we removed the labels from our GMap visualization. We then re-ran the entire session involving the memorability task, in conditions identical to our initial studies, on this modified GMap. Without the labels, the retrieval percentage dropped from approximately 0.23 to 0.1 which is about the same as the other three visualizations. To conclude, adding labels to the visualization improves memorability while other visual attributes have no noticeable impact.

D. An alternative GMap design

Based on our findings we hypothesized that GMap could outperform BubbleSets, especially in task 1, if separated areas belonging to the same group were linked explicitly, as in the case of BubbleSets or LineSets. To test this hypothesis, we manually altered the output of the GMap algorithm using image editing software to provide such links. An image is shown in Figure 6. This process of manually exploring a design is similar to the methodology proposed by the authors of LineSets. We then re-ran the first two tasks on this alternative design, in conditions identical to our initial studies. Results indicate that the modified GMap outperforms BubbleSets in fast membership assessment, as hypothesized, and provides about equal accuracy performance for deliberative assessment. (Fig. 6). This result provides additional support for the discussions in 4.2 and 4.3.

V. Discussion

An important number of errors and differences in our study were linked to assignment of similar colors to distinct sets. The BubbleSets and LineSets algorithms had a clear advantage over the other two visualizations because of their reduced reliance on color. Consequently, a legitimate question is whether increasing the color difference between clusters would render some of our results useless from a practical point of view.

We argue that the coloring approach is limited for two reasons. First, the more groups are present in a visualization, the harder it is to develop a set of distinct colors. A smart coloring algorithm similar to those described in [8, 28, 25, 24] could perhaps be developed to automatically choose perceptually different colors based on group proximities. Even such approaches however would be limited by the number of groups and the intricacy of the spatial layout.

Second, being able to use a limited range of color to display group information has certain advantages. For instance, GMap and BubbleSets use background color to describe groups, which leaves node color available to encode other data attributes, assuming the effect of color interactions is well understood. Other special considerations such as color blindness, integration into other systems and color schemes, or aesthetic appeal would also be easier to accommodate.

We also note that the results from the evaluated group tasks are likely transferable to any visualization for viewing group information event though they were analyzed in the context of node-link diagrams.

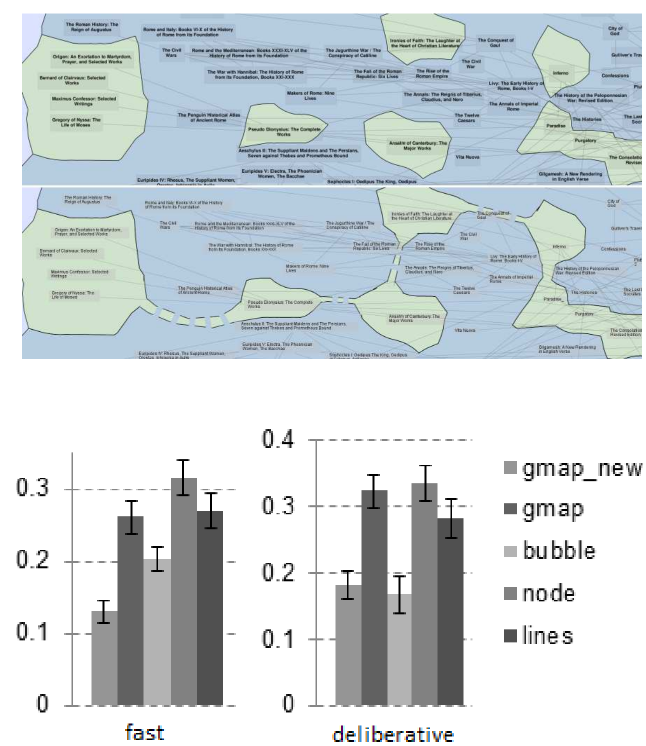


Figure 6: An improved design of Gmap that link disjointed set areas (top) performs equally or better than all other techniques.

Finally, we leave a few questions unaddressed. First, how do the observed differences change if color resolution is increased? This would allow us to isolate and understand the effectiveness of contours or space partitioning.

Second, BubbleSets gives control over how tightly contours are fitted around nodes, thus creating a continuum between space filling methods and individual nodes. It would be interesting to quantify the effect of this parameter change on membership assessment tasks.

Third, how would interaction play into our findings and what would be its impact?

Fourth, we have tested the four visualizations against each other on a single dataset and network layout. We believe further differences between techniques would be observed if the aspects of the underlying data and network were varied. Such aspects could include number and size of groups, spatial separation of groups, or their overlap.

Finally, we did not investigate the problem of overlapping groups. However, of the four tested visualizations only BubbleSets and LineSets enable the visualization of overlapping sets. These two techniques have already been compared in terms of set membership assessment, including tasks involving overlap perception, in the LineSets paper.

VI. Conclusion

We presented the results of a user study which combined online experimentation and eye-tracking analysis to compare four visualization methods for meshing group information and node link diagrams: GMap, BubbleSets, LineSets, and colored nodes. First, we found that BubbleSets is the best alternative for four types of tasks that involve group perception and understanding, with LineSets a viable alternative. Specifically, BubbleSets outperformed colored nodes in three of the four task categories, outperformed GMap in two of the four categories, and outperformed LineSets in just one of the four categories. Second, we found that visually encoding group information degrades performance on path tracing tasks. Third, we found that GMap's use of group labels improves group layout memorability. Finally, we discussed visual attributes of the four visualizations and their likelihood to explain the observed differences in performance.

VII. References

[1] Basak Alper, Nathalie Riche, Gonzalo Ramos, and Mary Czerwinski. Design study of linesets, a novel set visualization technique. Visualization and Computer Graphics, IEEE Transactions on, 17(12):2259–2267, 2011.

[2] Daniel Archambault and Helen C Purchase. The mental map and memorability in dynamic graphs. In Pacific Visualization Symposium (PacificVis), 2012 IEEE, pages 89–96. IEEE, 2012.

[3] Daniel Archambault and Helen C Purchase. Mental map preservation helps user orientation in dynamic graphs. In Graph Drawing, pages 475–486. Springer, 2013.

[4] A. Barsky, T. Munzner, J. Gardy, and R. Kincaid. Cerebral: Visualizing multiple experimental conditions on a graph with biological context. Visualization and Computer Graphics, IEEE Transactions on, 14(6):1253–1260, 2008.

[5] Florence Benoy and Peter Rodgers. Evaluating the comprehension of euler diagrams. In Information Visualization, *2007. IV'07. 11th International Conference*, pages 771–780. IEEE, 2007.

[6] Andrew Blake, Gem Stapleton, Peter Rodgers, Liz Cheek, and John Howse. Does the orientation of an euler diagram affect user comprehension? In DMS, pages 185–190, 2012.

[7] Christopher Collins, Gerald Penn, and Sheelagh Carpendale. Bubble sets: Revealing set relations with isocontours over existing visualizations. Visualization and Computer Graphics, IEEE Transactions on, 15(6):1009–1016, 2009.

[8] Çagatay Demiralp, Song Zhang, David F Tate, Stephen Correia, and David H Laidlaw. Connectivity-aware sectional visualization of 3d dti volumes using perceptual flat-torus coloring and edge rendering. In Proceeding of Eurographics 2006. Eurographics Association 2006.

[9] Andrew T Duchowski. A breadth-first survey of eye-tracking applications. Behavior Research Methods, Instruments, & Computers, 34(4):455–470, 2002.

[10] T. Dwyer, K. Marriott, F. Schreiber, P. Stuckey, M. Woodward, and M. Wybrow. Exploration of networks using overview+ detail with constraint-based cooperative layout. Visualization and Computer Graphics, IEEE Transactions on, 14(6):1293–1300, 2008.

[11] Michael Farrugia and Aaron Quigley. Effective temporal graph layout: A comparative study of animation versus static display methods. Information Visualization, 10(1):47–64, 2011.

[12] Emden R Gansner, Yifan Hu, and Stephen G Kobourov. Visualizing graphs and clusters as maps. Computer Graphics and Applications, IEEE, 30(6):54–66, 2010.

[13] Emden R Gansner and Stephen C North. An open graph visualization system and its applications to software engineering. Software Practice and Experience, 30(11):1203–1233, 2000.

[14] E.R. Gansner, Y. Hu, and S. Kobourov. Gmap: Visualizing graphs and clusters as maps. In Pacific Visualization Symposium (PacificVis), 2010 IEEE, pages 201–208. IEEE, 2010.

[15] Sohaib Ghani, Niklas Elmqvist, and Ji Soo Yi. Perception of animated node-link diagrams for dynamic graphs. In Computer Graphics Forum, volume 31, pages 1205–1214. Wiley Online Library, 2012.

[16] Steve Haroz and David Whitney. How capacity limits of attention influence information visualization effectiveness. Visualization and Computer Graphics, IEEE Transactions on, 18(12):2402–2410, 2012.

[17] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In Proceedings of the 28th international conference on Human factors in computing systems, pages 203–212. ACM, 2010.

[18] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on, pages 32–39. IEEE, 2005.

[19] Christian Heine and Gerik Scheuermann. Manual clustering refinement using interaction with blobs. In Proceedings of the 9th Joint Eurographics/IEEE VGTC conference on Visualization, pages 59–66. Eurographics Association, 2007.

[20] W. Huang and P. Eades. How people read graphs. In proceedings of the 2005 Asia-Pacific symposium on Information visualisation-Volume 45, pages 51–58. Australian Computer Society, Inc., 2005.

[21] W. Huang, P. Eades, and S.H. Hong. Beyond time and error: a cognitive approach to the evaluation of graph drawings. In Proceedings of the 2008 conference on BEyond time and errors: novel evaLuation methods for Information Visualization, page 3. ACM, 2008.

[22] W. Huang, P. Eades, and S.H. Hong. Measuring effectiveness of graph visualizations: A cognitive load perspective. Information Visualization, 8(3):139–152, 2009.

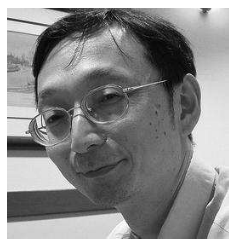
[23] W. Huang, S.H. Hong, and P. Eades. Predicting graph reading performance: a cognitive approach. In Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60, pages 207–216. Australian Computer Society, Inc., 2006.

[24] Christophe Hurter, Mathieu Serrurier, Roland Alonso, Gilles Tabart, and Jean-Luc Vinot. An automatic generation of schematic maps to display flight routes for air traffic controllers: structure and color optimization. In Proceedings of the

International Conference on Advanced Visual Interfaces, pages 233–240. ACM, 2010.
[25] R. Jianu, C. Demiralp, and D.H. Laidlaw. Exploring 3d dti fiber tracts with linked 2d representations. Visualization and Computer Graphics, IEEE Transactions on, 15(6):1449–1456, 2009.
[26] R. Jianu, C. Demiralp, and D.H. Laidlaw. Exploring brain connectivity with two-dimensional neural maps. Visualization and Computer Graphics, IEEE Transactions on, 18(6):978–987, 2012.
[27] R. Jianu and D.H. Laidlaw. Poster: Visualizing protein interaction networks as google maps. In Proceedings of InfoVis 2010, Poster Track, 2006.
[28] R. Jianu, A. Rusu, A.J. Fabian, and D.H. Laidlaw. A coloring solution to the edge crossing problem. In Information Visualisation, 2009 13th International Conference, pages 691–696. IEEE, 2009.
[29] R. Jianu, K. Yu, L. Cao, V. Nguyen, A.R. Salomon, and D.H. Laidlaw. Visual integration of quantitative proteomic data, pathways, and protein interactions. Visualization and Computer Graphics, IEEE Transactions on, 16(4):609–620, 2010.
[30] Jeff Johnson et al. Designing with the mind in mind: Simple guide to understanding user interface design rules. Morgan Kaufmann, 2010.
[31] Nicholas Kong, Jeffrey Heer, and Maneesh Agrawala. Perceptual guidelines for creating rectangular treemaps. Visualization and Computer Graphics, IEEE Transactions on, 16(6):990–998, 2010.
[32] B. Lee, C. Plaisant, C.S. Parr, J.D. Fekete, and N. Henry. Task taxonomy for graph visualization. In Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization, pages 1–5. ACM, 2006.
[33] Kim Marriott, Helen Purchase, Michael Wybrow, and Cagatay Goncu. Memorability of visual features in network diagrams. Visualization and Computer Graphics, IEEE Transactions on, 18(12):2477–2485, 2012.
[34] Luana Micallef, Pierre Dragicevic, and Jean-Daniel Fekete. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. Visualization and Computer Graphics, IEEE Transactions on, 18(12):2536–2545, 2012.
[35] H. Purchase. Which aesthetic has the greatest effect on human understanding? In Graph Drawing, pages 248–261. Springer, 1997.
[36] H. Purchase, R. Cohen, and M. James. Validating graph drawing aesthetics. In Graph Drawing, pages 435–446. Springer, 1996.
[37] H.C. Purchase, D. Carrington, and J.A. Allder. Empirical evaluation of aesthetics-based graph layout. Empirical Software Engineering, 7(3):233–255, 2002.
[38] Nathalie Henry Riche and Tim Dwyer. Untangling euler diagrams. Visualization and Computer Graphics, IEEE Transactions on, 16(6):1090–1099, 2010.
[39] Peter Rodgers, Leishi Zhang, and Helen Purchase. Wellformedness properties in euler diagrams: which should be used? Visualization and Computer Graphics, IEEE Transactions on, 18(7):1089–1100, 2012.
[40] Joel Ross, Lilly Irani, M Silberman, Andrew Zaldivar, and Bill Tomlinson. Who are the crowdworkers?: shifting demographics in mechanical turk. In *CHI'10 Extended Abstracts on Human Factors* in Computing Systems, pages 2863–2872. ACM, 2010.
[41] Paolo Simonetto, David Auber, and Daniel Archambault. Fully automatic visualisation of overlapping sets. In Computer Graphics Forum, volume 28, pages 967–974. Wiley Online Library, 2009.
[42] 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. In Computer graphics forum, volume 30, pages 1719–1749. Wiley Online Library, 2011.
[43] C. Ware, H. Purchase, L. Colpoys, and M. McGill. Cognitive measurements of graph aesthetics. Information Visualization, 1(2):103–110, 2002.
[44] Colin Ware. Toward a perceptual theory of flow visualization. Computer Graphics and Applications, IEEE, 28(2):6–11, 2008.
[45] Colin Ware and Glenn Franck. Evaluating stereo and motion cues for visualizing information nets in three dimensions. ACM Transactions on Graphics (TOG), 15(2):121–140, 1996.
[46] Colin Ware and Peter Mitchell. Visualizing graphs in three dimensions. ACM Transactions on Applied Perception (TAP), 5(1):2, 2008.

Radu Jianu received The PhD degree in computer science from Brown University. He is currently an Assistant Professor in the School of Computing and Information Sciences at Florida International University. His research interests are in visualization evaluation, applications of eye-tracking to data visualization, and applications of visualization to disaster management and mitigation.

Adrian Rusu is a Professor of Computer Science at Rowan University. He is the Director of the Software Engineering, Graphics, and Visualization (SEGV) research group. He received his Ph.D. in 2003 from University at Buffalo, SUNY.



Yifan Hu received the B.S. and M.S. in applied mathematics from Shanghai Jiao-Tong University (1981,1985) and Ph.D. in optimization from Loughborough University, UK (1992). He is a Principal Member of Technical Staff in the Information Visualization Department at AT&T Labs Research. He is a contributor to the Graphviz graph drawing system. His research interests include numerical and combinatorial algorithms, information visualization and data mining.

Douglas Taggart received his B.S. in computer science from Rowan University and is currently a research assistant at Rowan University. His interests are in visualization, mathematical modeling, and software engineering.