EuroRV3: EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (2017)
K. Lawonn, N. Smit, and D. Cunningham (Editors)

# Where'd it go? How geographic and force-directed layouts affect network task performance

Scott A. Hale[1,2], Graham McNeill[1], and Jonathan Bright[1]

[1]Oxford Internet Institute, University of Oxford, UK
[2]Alan Turing Institute, UK

**Abstract**

*When visualizing geospatial network data, it is possible to position nodes according to their geographic locations or to position nodes using standard network layout algorithms that ignore geographic location. Such data is increasingly common in interactive displays of Internet-connected sensor data, but network layouts that ignore geographic location data are rarely employed. We conduct a user experiment to compare the effects of geographic and force-directed network layouts on three common network tasks: locating a node, determining the path length between two nodes, and comparing the degree of two nodes. We found a geographic layout was superior for locating a node but inferior for determining the path length between two nodes. The two layouts performed similarly when participants compared the degree of two nodes. We also tested a relaxed- or pseudo-geographic layout created with multidimensional scaling and found it performed as well or better than the pure geographic layout on all tasks but remained inferior to the force-directed layout for the path-length task. We suggest interactive displays of geospatial network data allow viewers to switch between geographic and force-directed layouts, although further research is needed to understand the extent to which viewers are able to choose the most appropriate layout for a given task.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation; H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: User Interfaces—Evaluation/methodology

## 1. Introduction

Geospatial network data is almost always visualized by positioning nodes according to their geographic locations. It is known, however, that pure geographic positioning can interfere with other aspects of readability and comprehension in a network [WPCM02]. Several studies have therefore sought to create relaxed- or pseudo-geographical layouts that tweak network visualizations to improve angular resolution [BST00, CR14], avoid edge crossings [CDR04], maintain relative distances [Cla77, FMRM04], or optimize other metrics [PCA02].

Our study adds to this literature by comparing different network tasks on geographic and force-directed layouts using a controlled experiment with over 200 participants. Building on literature that compared multiple force-directed layouts and found that the best layout of a network depends on the task to be performed [DLF*09], we find that a geographic layout was superior for locating a node in our experiment, but that a force-directed layout was superior when determining the path length between two nodes. Our results showed no difference in the performance of people comparing the degree of two nodes. Given the increasing use of interactive visualizations for this type of data, we suggest that allowing users to switch between geographic and force-directed layouts (as well as pseudo-geographical layouts in-between these two endpoints) is an important option.

## 2. Experimental design

Our experiment involved observing both accuracy and completion time for common network tasks [LPP*06] with three different network visualizations.

### 2.1. Tasks

Various tasks have been used in the literature to examine different network layouts [LPP*06, DLF*09, PSD09, PCA02]. Using the topology-based tasks from the taxonomy developed by Lee et al. as a guide [LPP*06], we selected three popular tasks from the literature that would require participants to focus on local, regional and global network properties.

For our local scope task, we used an adjacency task, namely node degree [XRP*12, APP11], asking participants to compare the degree of two highlighted nodes. For our regional task, we used a connectivity tasks, namely path length [PSD09, HvW09], asking participants to determine the length of the shortest path between two highlighted nodes [WPCM02]. Finally, for our global





DOI: 10.2312/eurorv3.20171108

task we asked participants to locate a specified node within a network [PSD09], which Lee et al. note is "a common starting point for many tasks" [LPP*06].

### 2.2. Visualizations

We selected three different types of visualizations for the experiment. First, a simple geographic layout, where nodes were positioned according to their actual geographic locations. Such a geographic layout should best preserve the viewer's mental model of geography allowing them to draw upon existing knowledge to perform tasks.

Second, we chose a force-directed layout, the most common option for networks without geographic information [Kre11]. In particular, we used the Fruchterman-Reingold algorithm [FR91], which places nodes sharing edges near each other and the most well-connected nodes near the center of the network. With a force-directed layout, we expect regional tasks involving a small subset of nodes (e.g., path length) to be easier to accomplish. This follows from the notion that all the nodes involved should be positioned close to one another, allowing the participant to easily focus on the subpart of the network of interest and ignore the remainder of the network. We also note that with geographic positioning viewers often consider the straight line distance between points rather than the network/path-length distance [FMRM04, FM08], which again suggests force-directed layouts will be better for network distance (i.e., path-length) tasks.

Finally, given the hypothesized advantages of both geographic and force-directed layouts, we include a simple "pseudo-geographical layout." Our main focus is on comparing geographic and force-directed layouts, but many studies have shown a small sacrifice in geographic accuracy can improve visual analysis [BST00, CR14, CDR04, Cla77, FMRM04, PCA02]. Thus, our pseudo-geographic layout, which we design to avoid overly dense clusters while maintaining relative geographic positions, should perform better than a pure geographic layout and provide a stronger, more realistic comparison for the force-directed layout. Our approach using multidimensional scaling is detailed in the supplemental information.

### 2.3. Hypotheses

Based on the literature, we specifically tested the following hypotheses:

**H1** participants within our experiment will be able to locate a specific node within the network more accurately and more quickly when using the geographic layout.

**H2** participants within our experiment will be able to determine the path length between two nodes more accurately and more quickly when using a force-directed layout.

**H3** participants within our experiment will be able to compare the degree of two nodes more accurately and more quickly when using a force-directed layout.

## 3. Experimental procedure

The experiment was divided into three sections, each reflecting one of the three tasks attempted (node search, degree and path length). Each participant completed each of the three sections, and the order they received the sections in was randomized. Prior to starting each task, each participant was given a concise introduction to the requisite concepts and terminology. They then received two batches of three questions: three pretest questions assessing general background knowledge and three evaluated task questions.

Each participant was assigned to a single layout type at random (i.e., force-directed, geographical, or pseudo-geographical), and we used the same layout for the participant throughout all main sections. All layouts were computed in advance of the experiment; thus, although the Fruchterman-Reingold force-directed layout algorithm is non-deterministic all subjects in the force-directed condition saw the same realization of the layout. In all conditions, the label for a given node (the name of the local authority) was shown when the participant hovered the mouse over the node. When the mouse was not over a node, no labels were shown due to space limitations.

For the data underlying our experiments, we used a network of popular rail commuting routes between local authorities (administrative regions) within the United Kingdom. The network had 347 nodes (one for each local authority with all London local authorities grouped to one node) and 383 edges (Figure 1). For the degree and distance tasks, the participant was asked three pretest questions about a simple artificial network, followed by three questions about the actual network of commuting routes. For the node search task, the participant was asked as a pretest to find three local authorities on a simple polygon map, and then asked to find three further local authorities within the visualized commuting network. The order of the map and network for node search task were randomized to check for order effects.

The experiment was delivered using the academic crowdsourcing platform Prolific.ac. We restricted the study to people resident in the UK with a Prolific.ac approval rating of at least 90%. Source code for the experiment is available at: `http://www.github.com/oii-nexus/qa`. We rewarded each participant with a base, show-up fee of £0.50 as well as a bonus of £0.075 per correct answer (rounded up where necessary). Thus, the maximum total possible payout per participant was £1.85 $(0.50 + 18 \times 0.075)$. The per-question bonus ensured that participants were incentivized to answer all questions as accurately as possible. We forced participants to spend at least 10 seconds per question (to avoid undue rushing), but also prevented participants from spending longer than 30 seconds per question (to ensure all participants would complete the experiment in a reasonable time and ensure our payment resulted in an ethical wage). The time remaining was shown with a progress bar.

## 4. Results

201 participants completed the full experiment. We did not analyse responses from six further participants, three of whom experienced data loss and three of whom left the experiment early. Of the 201 participants with complete responses, 68 were randomly assigned to the force-directed condition, 70 to the geographic condition, and 63 to the pseudo-geographic condition. All participants were resident in the UK with 90% born in the UK and 89% having

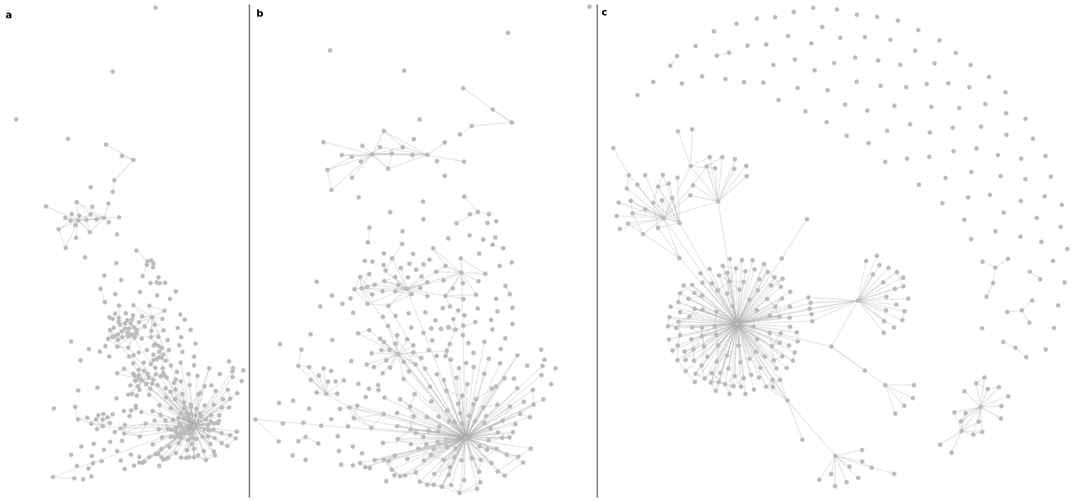


**Figure 1:** *Network of rail commuting routes between UK local authorities. (a) pure geographic layout; (b) MDS $\lambda = 150$ km; (c) force-directed.*

UK nationality. Most participants therefore had basic familiar with UK geography, but were unlikely to be familiar with the particular network data we used as it presented popular commutes by rail rather than physical rail lines. The mean age of participants was 34 (standard deviation 12), 57% of participants were female and 87% of participants classified themselves as Caucasian or white with the remaining 13% spread over 7 different ethnic groups. 56% of participants had an undergraduate degree or higher. Participants spent, on average, 6 minutes answering all pretest and main questions (standard deviation 1.9 minutes). The average time taken for a single question was 12.5 seconds (standard deviation 7.9 seconds).

### 4.1. Accuracy

The average participant answered 7 of the 9 main questions correctly (78%, standard deviation 1.79 questions). Participants had the highest accuracy with the degree task (average of 90% correct) and the lowest accuracy with the path length task (average of 56% correct). Participants answered a mean 64% of node search questions correctly. Not providing an answer within the 30 second limt was treated as an incorrect answer.

Figure 2 shows a breakdown of accuracy by condition. We tested for statistical significance using Welch two sample t-tests and employed a Bonferroni correction so that we only accepted p-values less than 0.017 (0.05/3) as significant. For the node search task, participants were significantly more accurate within both the pseudo-geographic condition (mean 2.17 questions, 72% correct, $p = 0.01$) and the geographic condition (2.03 questions, 68% correct, $p = 0.001$) than the force-directed condition (mean 1.57 questions, 52% correct). These results support hypothesis H1 that the geographic layouts would perform better than the force-directed layout for the node search task. The difference between the pseudo-geographical and geographical layouts is not significant ($p = 0.40$).

For the path length task, H2 hypothesized the force-directed condition would be most accurate, and the data support this. Participants were significantly more accurate within the force-directed condition (mean 2.22 questions, 74% correct) than either the geographic ($p < 10^{-6}$) or pseudo-geographic ($p < 10^{-7}$) condition. Participants within the geographic condition answered a mean 1.37 questions (46%) correctly, while participants within the pseudo-geographic condition answered a mean 1.43 questions (48%) correctly. Once again the difference between the geographic and pseudo-geographic conditions is not significant ($p = 0.71$). Finally in contrast to the prediction of H3, there are no significant differences between the accuracies of participants on the degree task within any of the three conditions.

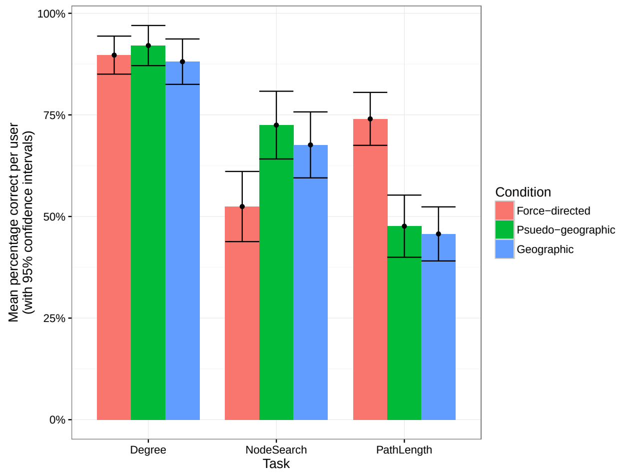


**Figure 2:** *The percentage of overall questions answered correctly by condition with 95% confidence intervals.*

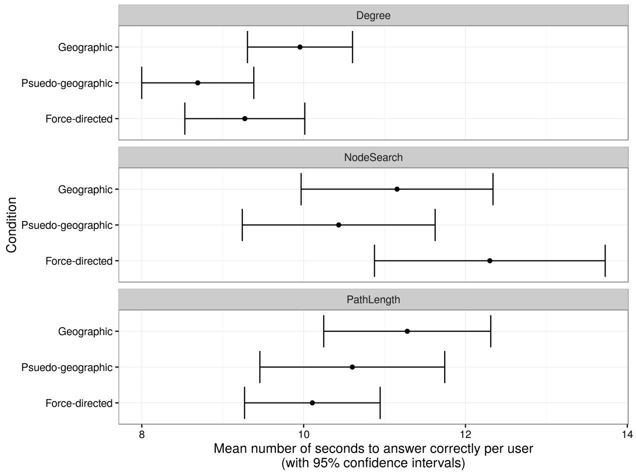


**Figure 3:** *The mean amount of time in seconds spent to answer correctly by task and condition with 95% confidence intervals.*

### 4.2. Time spent

In addition to accuracy, we measured the total number of seconds that participants spent answering each question in each section. The mean number of seconds to answer correctly for each task and condition are shown in Figure 3. We performed similar statistical analysis using Welch two sample t-tests with a Bonferroni correction and observed few statistical differences between the three conditions. This is most likely due to our limit of 30 seconds to answer each question. Some participants who were unable to answer correctly within the time limit may have answered correctly if they had been given more time. Due to the time limit, this difference is reflected in the accuracy scores already reported rather than the time spent per question. As mentioned earlier, we selected this upper time bound to avoid subjects spending too long on any one question and to ensure we paid an ethical wage to our participants for their time.

The only significant difference in the time to answer correctly is between the geographic and pseudo-geographic conditions in the degree task. There participants correctly answered questions significantly quicker within the pseudo-geographic condition (8.69 seconds) than within the geographic condition (9.95 seconds, $p = 0.009$). Neither of these values differed significantly from the 9.27 seconds taken to answer correctly within the force-directed condition.

## 5. Discussions

Our results indicate that there is a trade-off in performance and accuracy that depends both on the task to be completed and the layout of the graph. As geospatial network data from sensors and other new sources increases, designers of visual interfaces need to be mindful of this. Geographic layouts preserve a person's mental model and allow the person to complete tasks with global scope (e.g., locating a particular node) quickly and accurately. However, geographic layouts can make tasks involving a subset of nodes (e.g., determining the path length between two nodes) more difficult resulting in lower accuracies compared to force-directed layouts. Research indicates pseudo-geographic layouts can improve upon pure geographic layouts [BST00, CR14, CDR04, Cla77, FMRM04], and this study shows that unconstrained force-directed layouts are worth further investigation.

Once a person has located a given node of interest and is performing a relatively local task (e.g., determining the degree of a node) we found no difference between the force-directed condition and either geographic layout condition. While the accuracies were similar, we found that participants using the pseudo-geographical layout were faster than participants using the geographic layout for the local, degree task.

Nodes in the pseudo-geographic layout were more spread out with fewer dense "clumps" (Figure 1), and this likely also improved angular resolution [BST00]. We only studied one network (where nodes were local authorities in the UK), and the extent to which node overlap/occlusion or simply high node density affects pure geographic layouts will obviously depend on the specific network being visualized.

## 6. Conclusions

This study is only a first step and further experiments with additional datasets and tasks are needed before generalizations can be made. Nonetheless, for our data and tasks participants were more accurate with an unconstrained force-directed layout when evaluating the path length between two nodes in this study. Designers of interactive data exploration environments should consider allowing users to temporarily ignore the geographic dimensions of data and employ force directed layouts when needed. Most experimental studies to date compare relaxed- or pseudo-graphical layouts to pure geographic layouts, but such layouts are rarely compared to unconstrained force-directed layouts. Our findings show such comparisons are important.

An obvious advantage of interactive layouts is that the user is in control and can generally change the way a network is displayed. Within our study this was restricted to zooming and panning, but more generally interfaces may allow nodes to be filtered or re-positioned. Our results indicate it can be useful to temporarily discard geographic data and use force-directed layouts for certain tasks. An exciting avenue for future research, however, is to understand whether people can determine the best layout for a particular task when given the option to switch between layouts. That is, if we had given participants the ability to chose between a geographic or a force-directed layout would they have chosen the optimal layout for each task? If not, how could an interface better guide people towards the best layout for a given task? These are important questions as the amount of geospatial network data increases and as the use of interactive data exploration environments increases.

## 7. Acknowledgments

This publication was supported by Innovate UK as part of the NEXUS project (grant reference NE/N00728X/1) and The Alan Turing Institute (EPSRC grant EP/N510129/1). We further thank all our participants and the staff at Prolific.ac.

## References


[APP11] Archambault D., Purchase H., Pinaud B.: Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Transactions on Visualization and Computer Graphics 17*, 4 (2011), 539–552. doi:10.1109/TVCG.2010.78. 1

[BST00] Brandes U., Shubina G., Tamassia R.: *Improving Angular Resolution in Visualizations of Geographic Networks*. Springer Vienna, Vienna, 2000, pp. 23–32. doi:10.1007/978-3-7091-6783-0_3. 1, 2, 4

[CDR04] Cabello S., Demaine E. D., Rote G.: *Planar Embeddings of Graphs with Specified Edge Lengths*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 283–294. doi:10.1007/978-3-540-24595-7_26. 1, 2, 4

[Cla77] Clark J. W.: Time-distance transformations of transportation networks. *Geographical Analysis 9*, 2 (1977), 195–205. doi:10.1111/j.1538-4632.1977.tb00573.x. 1, 2, 4

[CR14] Chivers D., Rodgers P.: *Octilinear Force-Directed Layout with Mental Map Preservation for Schematic Diagrams*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 1–8. has an experiment that compares very similar things to the experiment presented in this paper. doi:10.1007/978-3-662-44043-8_1. 1, 2, 4

[DLF*09] Dwyer T., Lee B., Fisher D., Quinn K. I., Isenberg P., Robertson G., North C.: A comparison of user-generated and automatic graph layouts. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (Nov 2009), 961–968. doi:10.1109/TVCG.2009.109. 1

[FM08] Fabrikant S. I., Montello D. R.: The effect of instructions on distance and similarity judgements in information spatializations. *International Journal of Geographical Information Science 22*, 4 (2008), 463–478. doi:10.1080/13658810701517096. 2

[FMRM04] Fabrikant S. I., Montello D. R., Ruocco M., Middleton R. S.: The distance–similarity metaphor in network-display spatializations. *Cartography and Geographic Information Science 31*, 4 (2004), 237–252. doi:10.1559/1523040042742402. 1, 2, 4

[FR91] Fruchterman T. M. J., Reingold E. M.: Graph drawing by force-directed placement. *Software: Practice and Experience 21*, 11 (1991), 1129–1164. doi:10.1002/spe.4380211102. 2

[HvW09] Holten D., van Wijk J. J.: A user study on visualizing directed edges in graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2009), CHI '09, ACM, pp. 2299–2308. URL: http://doi.acm.org/10.1145/1518701.1519054, doi:10.1145/1518701.1519054. 1

[Kre11] Krempel L.: Network visualization. In *The SAGE Handbook of Social Network Analysis*, Scott J., Carrington P. J., (Eds.). SAGE Publications, 2011. 2

[LPP*06] Lee B., Plaisant C., Parr C. S., Fekete J.-D., 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* (New York, NY, USA, 2006), BELIV '06, ACM, pp. 1–5. URL: http://doi.acm.org/10.1145/1168149.1168168, doi:10.1145/1168149.1168168. 1, 2

[PCA02] Purchase H. C., Carrington D., Allder J.-A.: Empirical evaluation of aesthetics-based graph layout. *Empirical Software Engineering 7*, 3 (2002), 233–255. doi:10.1023/A:1016344215610. 1, 2

[PSD09] Pohl M., Schmitt M., Diehl S.: Comparing the readability of graph layouts using eyetracking and task-oriented analysis. In *Computational Aesthetics* (2009), pp. 49–56. 1, 2

[WPCM02] Ware C., Purchase H., Colpoys L., McGill M.: Cognitive measurements of graph aesthetics. *Information Visualization 1*, 2 (2002), 103–110. doi:10.1057/palgrave.ivs.9500013. 1

[XRP*12] Xu K., Rooney C., Passmore P., Ham D. H., Nguyen P. H.: A user study on curved edges in graph visualization. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (Dec 2012), 2449–2456. doi:10.1109/TVCG.2012.189. 1