Investigating the Efficacy of Network Visualizations for Intelligence Tasks

Christopher W. Berardi, Erin T. Solovey, Mary L. Cummings
Humans and Automation Laboratory, Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract— There is an increasing requirement for advanced analytical methodologies to help military intelligence analysts cope with the growing amount of data they are saturated with on a daily basis. Specifically, within the context of terror network analysis, one of the largest problems is the transformation of raw tabular data into a visualization that is easily and effectively exploited by intelligence analysts. Currently, the primary method within the intelligence domain is the node-link visualization, which encodes data sets by depicting the ties between nodes as lines between objects in a plane. This method, although useful, has limitations when the size and complexity of data grows. The matrix offers an alternate perspective because the two dimensions of the matrix are arrayed as an actors x actors matrix. This paper describes an experiment investigating node-link and matrix visualization techniques within social network analysis, and their effectiveness for the intelligence tasks of: 1) identifying leaders and 2) identifying clusters. The sixty participants in the experiment were all Air Force intelligence analysts and we provide recommendations for building visualization tools for this specialized group of users.

Keywords—Terrorism; Terrorist Network; Visualization; Node-Link; Matrix; Intelligence Analyst; Social Network Analysis

I. INTRODUCTION

Military intelligence analysts are increasingly tasked to sift through enormous volumes of data to identify the proverbial intelligence "needle in a haystack." With the fast paced development cycle for new sensors, which can collect data at unmanageable rates, military intelligence analysts need analytical techniques that give them the ability to effectively and efficiently transform the collected data into intelligible information and, subsequently, intelligence. Specifically, within the context of terror network analysis, one of the largest problems is the transformation of raw tabular data into a visualization that is effectively exploited by analysts.

The current method within the intelligence domain is the node-link visualization, which encodes data sets by depicting the ties between nodes as lines between objects in a plane. This method has limitations when the size and complexity of data grows [1]. Therefore, this research was motivated by the desire to evaluate the matrix visualization using analysts to assess the efficacy of this form of visualization and potentially identify an alternate means of visualizing terror networks.

The matrix offers an alternate perspective because the two dimensions of the matrix are arrayed as an actors x actors matrix, which implies the same layout of actors contained on the rows is also contained on the columns. A relationship between actors is communicated by a Boolean value where the rows and columns of specific nodes intersect. This form of visualization may offer benefits over the node-link, because as node-link visualizations grow in size, they have a tendency to occlude data. Matrices offer a solution to his problem, because in matrix visualizations objects cannot overlap; thus resolving the data occlusion and improving readability.

The primary focus of this paper is to analyze two visualization techniques within social network analysis, with the intent to explore their effectiveness for the intelligence tasks of: 1) identifying leaders within a terror network and 2) identifying clusters within a terror network. We first provide background on the intelligence process and the role of network visualizations. Then, we describe the most common visualization method (node-link) as well as an alternative method (matrix). This is followed by details of the experiment that was conducted to assess the effectiveness of these methods in typical intelligence tasks as well as the results of the study. Finally, we discuss some of the implications and conclusions that can be drawn from the experiment results.

II. THE INTELLIGENCE PROCESS

Intelligence is only of value when it is available and contributes to, or shapes, a decision-making process by, "providing reasoned insight into future conditions or situations" [2]. This same axiom does not hold true for raw data. Therefore, the burden is on the intelligence analyst to transform raw data into intelligence. This transformative process begins with the collection of data from sensors. The first step is to process the raw data into a form intelligible by an analyst (see Fig. 1). Depending on the type of raw data, this step is either automated as in the production of an image from a camera, or requires an analyst, in limited cases, to transform the raw data into information such as language translation. In the context of social network analysis, this stage typically involves transforming the tabular raw data into a visualization, or series of visualizations. This specific transformative process (data → information) is also known within the intelligence community as processing and exploitation.

During the processing and exploitation phase, as shown in Fig. 1, an analyst most commonly transforms the data into a node-link visualization. Little to no emphasis is given to creating alternating modes of visualization that could result in a more effective transformation of data to information. Furthermore, there is little existing research into the effectiveness of one form of visualization over another in the domain of intelligence [3].

This research was funded in part by MIT Lincoln Laboratory and the NSF (grant #1136996 awarded to Computing Research Association for the CI Fellows project)

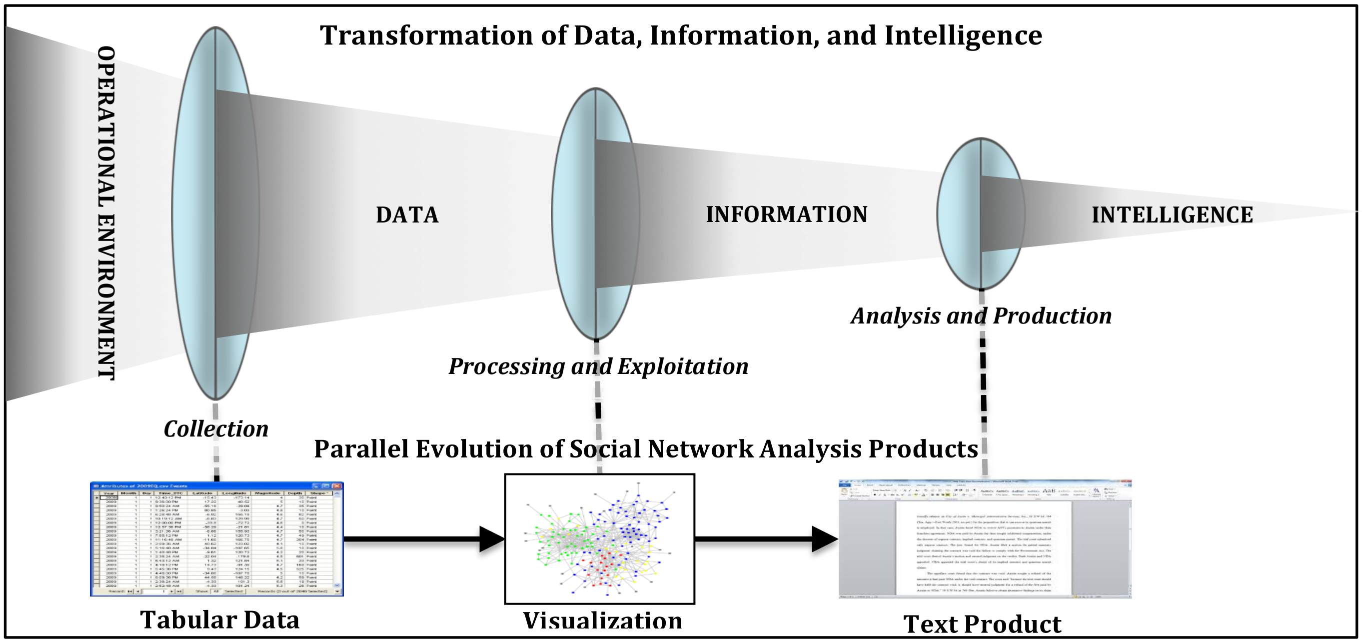

Figure 1: Parallel Evolution of Social Network Analysis Products

After data is transformed into information, the subsequent information can be integrated and analyzed to produce intelligence. During analysis, assessments are made by comparing already integrated and evaluated information; these assessments are combined and used to discern patterns or links. Finally, the analysis and production process concludes with interpretation, which is a largely inductive reasoning process in which available information is evaluated. From this sequence of integration, evaluation, analysis, and interpretation, intelligence is finally produced. Although this is a generic process which applies to all forms of intelligence, within the context of social network analysis, analysis this would be conducted by evaluating multiple visualizations of social networks and interpreting the information resident in each of those visualizations to create a prediction about the terror network, or networks, being analyzed (Fig. 1).

III. VISUALIZATION METHODS

There are numerous visualization methods for social network analysis. This work focuses on node-link and matrix visualizations as they have the most promise for intelligence tasks. For the most part, network visualizations seek to discover two types of patterns: 1) social groups, defined as collections of actors who are tightly linked to one another; or 2) social positions, defined as actors who are linked into a total social system in a defining way [4]. Quantitative measures have been developed to identify actors or groups of interest in social networks, and the most common are *betweenness centrality* (ability to connect groups) and *closeness centrality* (high degree of access to network). While these measures are unable to solely identify a leader within a network, it can be used to cue an analyst's attention on a specific actor for further study or analysis with other social network measures of centrality or ego.

A. Node-Link Visualizations

Within the military intelligence enterprise, node-link visualizations are by far the most pervasive for social network analysis. Research indicates that node-link diagrams are one of the most effective ways to visualize relatively low density networks (a link density less than 0.4) [1]. However, when the data density grows larger and more dimensions of data (such as node attributes) are added, node-link visualizations can become complex and difficult to analyze.

Fig 2 shows a node-link visualization organized using a spring embedded layout [5]. The primary benefit of node-link visualization is the ability to view the entire network topology. However, at its most rudimentary form, node-link analysis shows only nodes and links. It is possible, by overlaying additional node attributes, to show multiple dimensions of meta-information using this visualization technique. Fig. 2 demonstrates one method where meta-information can be incorporated into a node-link visualization. In this example, ordinal values of closeness centrality are depicted by changing node size; where a large node indicates high closeness centrality and a small node indicates low closeness centrality. Additionally, the scalar values are denoted on the left side of the label for each node.

Another dimension of data can be added by changing the node color. Fig 2 illustrates an example of this method. In this case, closeness centrality is still depicted by node size, but betweenness centrality is now illustrated by changing the node color; where blue nodes indicate high betweenness centrality and red nodes indicate low betweenness centrality. The result is a visualization that shows the topology, but also offers quantitative measure of centrality that communicates the relative "importance" of each node.

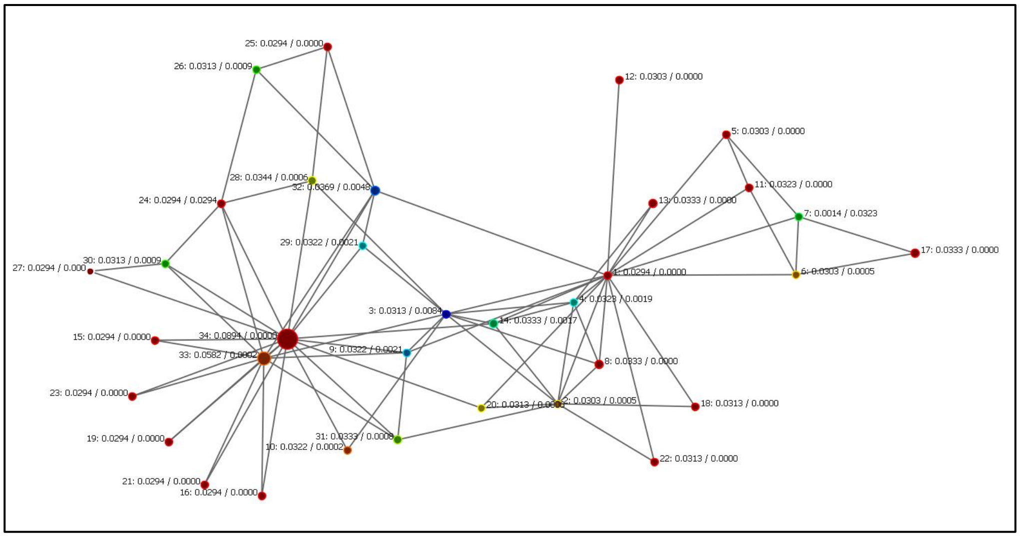

Figure 2: Enhanced node-link visualization with multidimensional data [10]

B. Matrix Visualizations

Matrix visualizations currently are not common within the intelligence community. The absence of this form of visualization may be due to a lack of commercially available intelligence oriented matrix analysis tools, and no analyst training on matrix analysis. This is supported in research by Henry et al., who surveyed all the tools in the "International Network for Social Network Analysis" software repository. Their examination revealed that node-link represented the preponderance (54 out of 55) of available tools [6].

Fig. 3 represents a matrix showing the interconnections between the nodes (labeled numerically from 1-34 on the top and left of the matrix). The matrix is symmetrical because the connections in this data set are nondirectional. Furthermore, in this matrix the links between nodes are represented by a "1" at the intersection of nodes (note all cells are with a "1" are highlighted in grey to make the clusters more apparent).

As with node-link visualizations, basic matrices only depict the topology. However, additional information may be encoded in the color and numerical value of the diagonal in the matrix.

IV. EXPERIMENT

The goal of this research was to investigate the two visualization methodologies to support intelligence analysis.

A. Participants

The 60 participants were all Air Force Airmen, with an average age of 32.92 ($\sigma = 9.59$), and who hold the Air Force Specialty Code of Intelligence Analyst.

B. Experimental Scenarios

Two experimental scenarios were used in the design of this experiment: 1) identifying leaders and 2) identifying clusters. These two tasks were chosen because of their importance highlighted in a literature review and a cognitive task analysis conducted prior to this experiment [3]. Additionally, this is corroborated by academic research into social network task taxonomy [7, 1] and recognized to be consistent with the primary tasks of social network analysis [8, 9].

C. Experimental Hypotheses

The ability to identify leaders from a node-link or matrix visualization is primarily influenced by the emergent features of the visualization techniques, such as the color proximity presented by conditionally formatting the measures of centrality and node position relative to the network topology. Node-link visualization may make this task easier by providing information on a node's position relative to the entire network topology. The matrix visualization also provides topological information, but to a lesser degree. Consequently, the following hypotheses capture the expected performance at identifying leaders:

- Hypothesis 1: The ability to accurately identify leaders within a network or cluster is expected to be better supported by the node-link visualization as compared to the matrix visualization.
- Hypothesis 2: Use of the node-link visualization is expected to require less time to accurately identify leaders as compared to the matrix visualization.

Successfully identifying clusters is primarily a function of two sub-tasks: 1) the ability to identify a cluster, and 2) to accurately identify where the cluster stops and where the next cluster begins (i.e. boundaries). The node-link visualization may make the second sub-task difficult, because there is no clear bifurcation between clusters. The matrix visualization may make this task easier, because it more clearly communicates the boundaries between clusters. However, the accuracy of the clusters within the matrix depends on the quality of the clustering algorithm used to partition the network. This factor and analyst unfamiliarity with this visualization technique may degrade matrix performance; however, the tutorial should help offset the second factor. Accordingly, the following hypotheses capture the expected performance at identifying clusters:

- Hypothesis 3: The ability to accurately identify clusters within a network is expected to be better supported by the

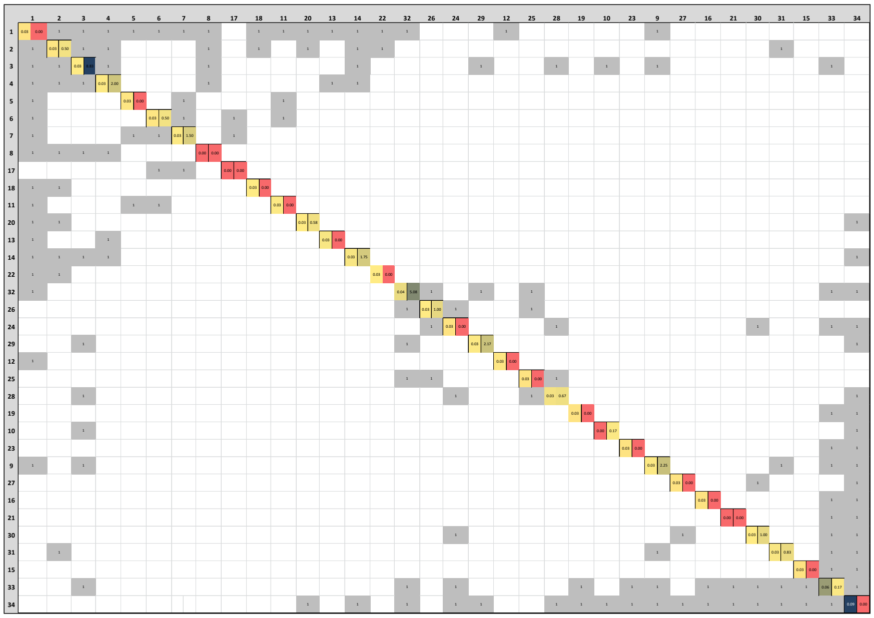

Figure 3: Partitioned symmetrical matrix of social network.

use of the matrix visualization as compared to the node-link visualization.

- Hypothesis 4: Use of the matrix visualization is expected to require less time to accurately identify leaders as compared to the node-link visualization.

D. Experimental Design and Tasks

To test the effectiveness of the visualizations, a 2 (Visualization Technique) x 2 (Visualization Task) mixed design study within-subjects on the visualization task factor and between-subjects on the visualization technique factor experiment was conducted in which participants exploited matrix or node-link visualizations constructed from a surrogate terror data set [10]. Each participant was given one visualization and asked to accomplish two tasks:

- (Task 1) "Analyze roles and positions - these are higher level tasks relying on the interpretation of groups of actors (positions) and connection patterns (roles)." [1]
    - (Question 1.1) Identify any central actors, which are defined as actors linked to many others or that bridge communities together.
    - (Question 1.2) Identify any potential leaders within the network.
    - (Question 1.3) Assuming there are only two leaders, identify those leaders.
- (Task 2) "Identify all communities, i.e. cohesive groups of actors that are strongly connected to each other." [1]
    - (Question 2.1) Assuming there are only two clusters, identify those clusters.

E. Independent Variables

Two independent variables were of interest in this experiment: 1) visualization technique and 2) visualization task. Visualization technique refers to specific visualization (node-link or matrix) a participant will use to answer the task questions. In this experiment participants saw either the node-link (Fig. 2) or the matrix (Fig. 3) visualization. Therefore, this two level factor was a between-subjects variable. The visualization task factor includes both identifying leaders and clusters. This factor is a within-subjects variable, because every participant was asked to perform both the tasks.

F. Dependent Variables

Two dependent variables were used in the experiment: accuracy of analyst assessment and time to reach assessment.

G. Experimental Procedure

Each participant performed the experiment individually. After finishing a demographic survey, participants were given a self-paced tutorial that detailed the purpose of the experiment and explained the visualization. Participants were not offered any incentives for performance.

Included in the tutorials were a series of training lessons on the measures of centrality (closeness and betweenness). Participants were then given examples with the measures of centrality and asked to assess which nodes displayed the highest measure for both closeness and betweenness centrality. Only after a participant was able to correctly identify each measure of centrality could he or she proceed to the actual experiment. If a participant failed the training he or she received additional instruction until able to successfully demonstrate proficiency with the measures of centrality.

Participants then completed the two test scenarios described in the previous sections for the visualization; lasting approximately 10 minutes. Visualizations were individually printed on an 11x17 inch piece of plain white copy paper and given to the participant. In an effort to prevent a possible order effect, the order in which visualizations and task scenarios were presented was randomized. Prior to beginning the tests, participants were informed that their accuracy of assessment was recorded and that the time to reach assessment was being recorded. Once the experiment was completed, feedback was solicited about the experience through a post-experiment questionnaire.

V. Results

The results presented in this section are organized by dependent variable. For all reported results, α = 0.05 unless otherwise stated. For question 1.1, three participants who were given the matrix misread the question so those three responses were removed from the data set; resulting in an uneven number of responses for each condition (matrix = 27, node-link = 30).

A. *Accuracy of Analyst Assessment*

The t-test results indicate that node-link visualization had a statistically significant higher percentage correct for questions 1.1 and 1.2. However, the participant responses for question 1.3 were not normal; therefore, a Mann-Whitney U test was used to identify any statistically significant difference. The result of this test also indicates that node-link had a statistically significant higher percentage correct. The boxplot in Fig. 4 shows the median percentage correct, quartiles and extreme values (outside the whiskers, which show 5-95 percentiles) for each question.

B. *Time to Complete Performance*

Time to complete the task was recorded for all previously discussed questions. An unpaired two-tailed t-test was performed on the completion time for each question to determine if there was a statistically significant difference in the time required to complete for a given visualization.

The results indicate that there was not a statistically significant difference in completion time for any of the questions. Fig. 5 shows the boxplots of median time, quartiles and any extreme values (outside the whiskers, which shows 5-95 percentiles) for questions 1.1, 1.2, 1.3, and 2.1.

VI. Discussion

A. *Identification of Leaders Performance*

Identification of leaders was classified by the percentage of correct responses for questions 1.1, 1.2, and 1.3. The results indicate that for each of these questions, the node-link showed a statistically significant higher average percentage correct than the matrix. These results are consistent with hypothesis 1, which postulated that the ability to accurately identify leaders within a network or cluster is expected to be better supported by the node-link as compared to the matrix.

However, the results also indicate that for questions 1.1 and 1.2 the node-link took longer on average, although not statistically significant, to analyze than the matrix. These results are inconsistent with hypothesis 2, which theorized; use of the node-link visualization is expected to require less time to accurately identify leaders as compared to the matrix.

B. *Identification of Clusters Performance*

The results for the identification of clusters, classified by the percent correct response for question 2.1, were not as straightforward as those from the identification of leaders task. Node-link showed a higher average percentage of correctly identifying clusters than the matrix, although not statistically significant. These results are not consistent with hypothesis 3, which theorized that the ability to accurately identify clusters within a network is expected to be better supported by the use of the matrix as compared to the node-link. However, there was a higher amount of variability in the responses for the node-link than the matrix. These results possibly indicate that although the matrix lacked in accuracy over the node-link, it showed improved precision over the node-link.

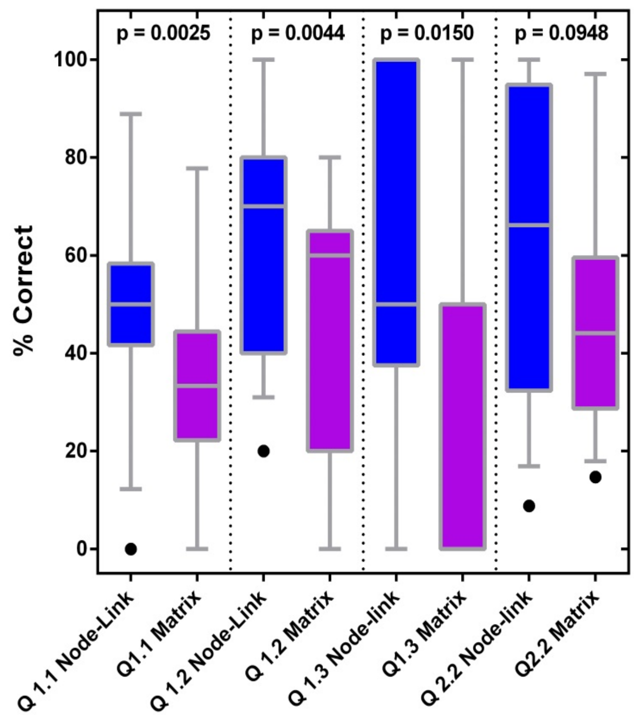

Figure 4: Percentage correct for questions 1.1, 1.2, 1.3, and 2.1.

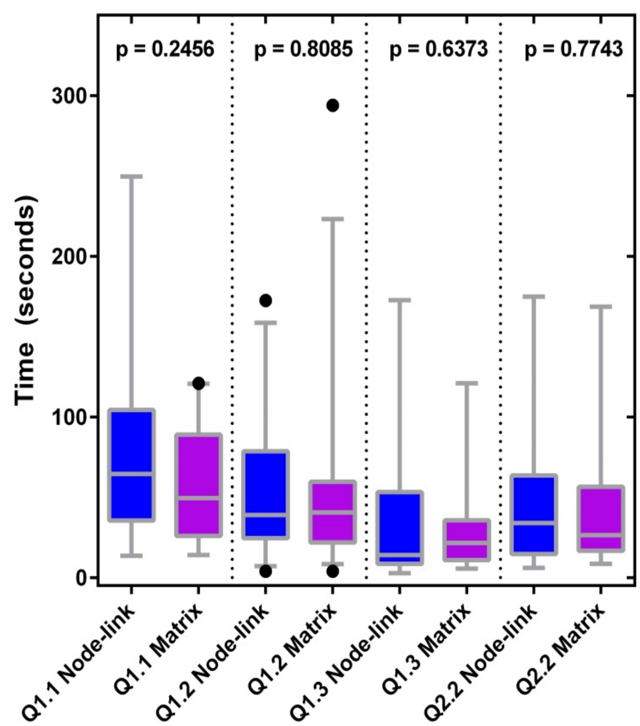

Figure 5: Completion time for questions 1.1, 1.2, 1.3 and 2.1.

Although the percentage correct favored the node-link, the matrix showed a shorter average time to complete, although not statistically significant, than the node-link. This result is consistent with hypothesis 4, which postulated that the use of the matrix visualization is expected to require less time to accurately identify leaders as compared to the node-link.

C. Subjective Responses

Many participants had difficulty understanding the numbers associated with the betweenness and closeness centrality. Roughly ten participants asked the investigator about the relative significance of the numbers and indicated the presence of the numbers complicated their analysis. Some ignored the numbers and relied on the colors and node size. The quick adjustment away from the numbers to the colors indicates that this section of the tutorial was effective and that the analysts generally understood how the measures of centrality could help them perform the tasks and employed gestalt-based reasoning.

Some participants felt recording the time forced them to rush through the experiment. Five participants even cited this as the component of the experiment they liked least, often indicating that it forced quick conclusions over thorough analysis. The utility offered by the collection of this variable may not be worth the impact on the overall experiment. This is a consideration which must be factored into future experiments.

Regardless of the visualization, the preponderance of participants recognized the importance and necessity of this type of research. Many either left feedback on the questionnaire similar to, "this is something we do not do enough of", or explicitly communicated this sentiment to the investigator after the experiment. While unsupported by any quantitative data, this resounding feedback indicates that analysts recognize that they may not be using the most effective means available to analyze terror networks. This conclusion does not support either visualization, but supports the need for continued research along the lines of this research.

VII. Recommendations and Future Work

Based on the quantitative assessment of the data collected and the comments elicited at the completion of the experiment, the following recommendations should be taken into consideration for future experiments into the effectiveness of visualizations for terror network analysis.

Continue to use intelligence analysts to test the effectiveness of intelligence visualizations. This demographic responded differently to the visualizations than hypothesized, in part because the hypotheses were based on research in academic literature where the participants were not intelligence analysts [1]. Initial conclusions supported by this paper indicate that the results of previous academic work on the effectiveness of social network visualizations may not be wholly extensible to the domain of intelligence.

Time to complete each task should not be an explicit component of the experiment. Although analysts often work under time pressure, the time pressure is not on the order of seconds or minutes. Furthermore, they were unaccustomed to having their analysis timed on a stopwatch. Removing this factor should help eliminate the effects resulting from analysts rushing through the analysis and potentially reaching premature conclusions as documented in the post experiment comments.

To gain a more detailed understanding of the potential of matrices, a *longitudinal study should be conducted using the matrix,* where a static group of analysts are tasked to exploit multiple terror networks at different times over a predetermined period of time; perhaps a month. This type of experiment would provide detailed data on learning that may occur as analysts become more familiar with the matrix.

While the node-link proved superior in the research outlined herein, this may be a byproduct of the specific data set chosen, not the visualization. To understand the extent of these affects, *more experiments are needed using a variety of data sets*. In addition, an experiment similar to the one described in this paper should be run for both the node-link and matrix *on a computer*. This form of experimentation is required to understand the effectiveness in a more representative setting.

VIII. Conclusion

Although the results of this work indicate that the node-link supported both tasks better than the matrix, more investigation is needed to determine if this conclusion is universal across all intelligence tasks and populations.

At this time, the matrix should not be universally integrated into the current methodologies used by analysts to exploit terror network visualizations until more research is conducted into the strengths and weaknesses within the intelligence domain. However, analysts should be independently encouraged to explore and adapt new methods of visualization into their current practices and identify new or improved versions of the visualizations identified within this paper for future testing.

References

[1] M. Ghoniem, J.-D. Fekete and P. Castagliola, "A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations," *Visualization and Computer Graphics, IEEE Transactions on,* vol. 13, no. 6, pp. 17-24, 2004.

[2] Joint Chiefs of Staff, Joint Intelligence, Vols. JP 2-0, 2012.

[3] C. W. Berardi, "Investigating the Efficacy of Terrorist Network Visualization," M.S. Thesis, MIT Aeronautics and Astronautics, Cambridge, MA, 2013.

[4] L. Freeman, "Visualizing Social Networks," *Journal of Social Structure,* vol. 1, no. 1, 2000.

[5] J. Barnes and P. Hut, "A hierarchical O(N log N) force-calculation algorithm," *Nature,* vol. 324, no. 4, pp. 446-449, 4 December 1986.

[6] N. Henry, J.-D. Feteke and M. J. McGuffin, "NodeTrix: A Hybrid Visualization of Social Networks," *IEEE Trans. on Visualization and Computer Graphics,* pp. 1302-1309, 2007.

[7] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete and N. Henry, "Task Taxonomy for Graph Visualizaiton," in *In Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization*, Venice, Italy, 2006.

[8] S. Wasserman and K. Faust, Social network analysis: methods and applications, New York: Cambridge University Press, 1994.

[9] J. P. Scott, Social Network Analysis: A Handbook, Sage Publications, 2000.

[10] W. W. Zachary, "An Information Flow Model for Conflict and Fission in Small Groups," *Journal of Anthropological Research,* vol. 33, no. 4, pp. 452-473, 1977.