TaMax: Visualizing Dense Multivariate Networks with Adjacency Matrices

Ilkin Safarli*
University of Utah

Alexander Lex†
University of Utah

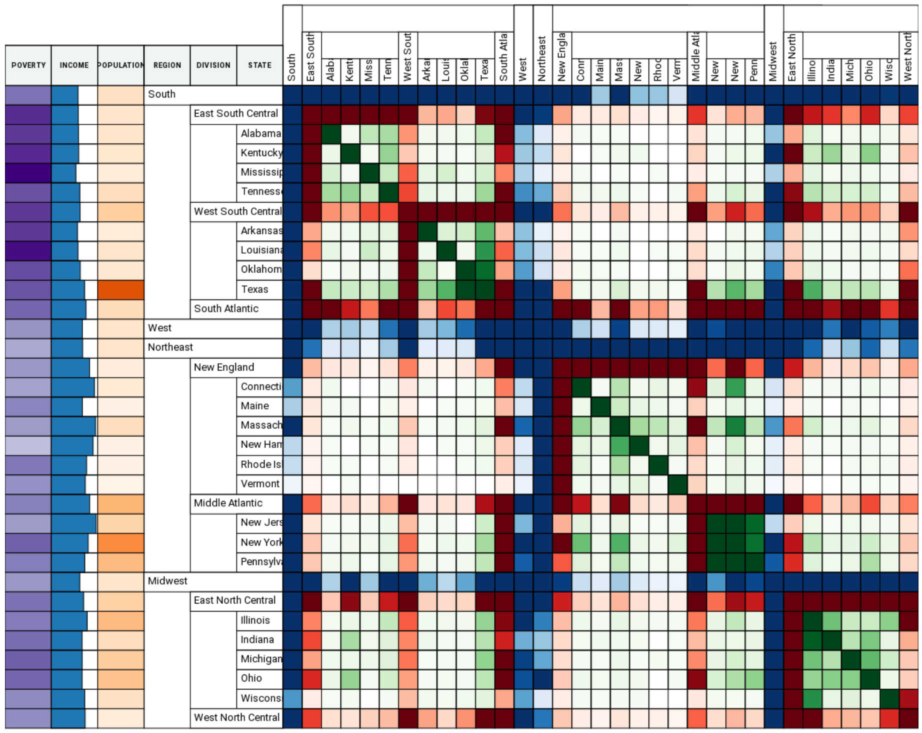


Figure 1: TaMax visualizing migrations between US states in 2016. The nodes are hierarchically grouped based on their division and region. Node attributes are shown in the node attribute table, and the migration flow is represented with an adjacency matrix. In the matrix, each aggregation level is represented using a different color scale.

Abstract

Considering node and edge attribute is crucial for many network exploration and analysis tasks. However, effective visualization of both structure and attributes is a challenging problem, especially for dense graphs. In this poster, we introduce TaMax, a technique designed to visualize dense multivariate graphs with a diverse set of node and edge attributes based on adjacency matrices. In TaMax, node attributes are visualized in a table that is juxtaposed with the matrix, while edge attributes visualized in the cells. We investigate different ways to visualize multiple edge attributes: dividing each cells into sub-cells showing different edge attributes or overlaying a secondary attribute with opacity over a cell. Furthermore, TaMax addresses the scalability problem by allowing flexible grouping based on node attributes and querying based on edge attributes.

Index Terms: Human-centered computing—Visualization—Visualization techniques

1 Introduction

Multivariate graphs have both nodes and links connecting these nodes, and attributes about the nodes and/or links. In many analysis scenarios, understanding these attributes is just as important as understanding the network data itself. However, visualizing attributes at the same time as the network structure is a challenging problem, and many methods have been developed for visualizing multivariate graphs [4] focusing on different network types and use cases. However, current techniques often either focus on visualizing node or edge attributes. Many approaches also are not well suited for dense multivariate graphs. In this poster, we present TaMax, a table+matrix approach for visualizing multivariate networks with a diverse set of node and edge attributes. TaMax facilitates visualizing multiple node and edge attributes, grouping based on node attributes, rearranging the matrix, and querying based on edge attributes.

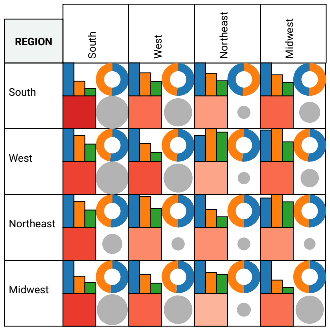
2 Related Work

None of the existing multivariate network visualization tools support dense networks with many attributes well. However, as Nobre et al. [4] discuss, matrices have untapped potential for visualizing this type of networks, hence, we will only discuss adjacency matrix based methods here. For a broader overview, we refer to the review by Nobre et al [4].

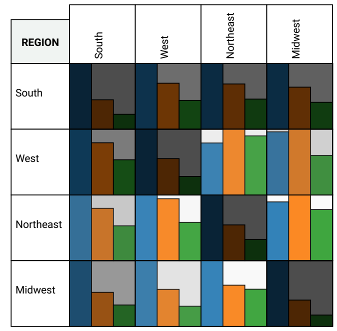
ZAME [2] is a technique that is designed for visualizing and exploring large multivariate networks using adjacency matrices. Its key contribution is scalability, but ZAME can also visualize distributions of attributes for aggregated nodes. While ZAME supports aggregation, the aggregation method cannot be specified by the user.

Alper et al. [1] propose a similar method to ZAME, where different encodings are used to visualize edge attributes in the matrix cells. The use case for multiple attributes in Alper et al.'s work is graph comparison. Both ZAME and the technique proposed by Alper et al. do not provide support for visualizing node attributes.

*e-mail: ilkin@sci.utah.edu
†e-mail: alex@sci.utah.edu



(a) Multiple edge attributes.



(b) Edge counts overlaid with the cells.

Figure 2: Different methods for visualizing multiple edge attributes. (a) The **cell division approach** shows multiple attributes of an edge in sub-cells. The histogram shows the education level of migrants, the doughnut chart shows male to female ratio, average age is encoded with the saturation of the red rectangles, and the area of the circles represent average income of the people who are migrating. (b) The **overlay approach** shows a base variable, a histogram of education levels in this case, overlaid with a transparent rectangle that encodes a numerical attribute using opacity; the total number of migrants in this case.

Graffinity [3] uses a matrix to visualize higher-level connectivity, as opposed to adjacency. It supports attribute-driven aggregation, and visualizes attributes in a juxtaposed table. However, Graffinity does not have support for multiple edge attributes.

3 TaMax Design

TaMax is designed for visualizing dense or even completely connected networks with many and diverse node and edge attributes. We chose an adjacency matrix as our basic network encoding, as it is well suited to visualize dense networks and can visualize edge attributes well. To make TaMax scalable, we use a variety of interaction methods, including dynamic grouping, filtering, and re-ordering.

As shown in Fig. 1, TaMax has two main views, an adjacency matrix view and a juxtaposed table view. Supplementary views are used to select which attributes to display, and query the network based on the attributes.

Matrix View. The adjacency matrix visualize the network and edge attributes. Matrices are commonly used to visualize edges and edge weights, but rarely to also visualize additional attributes. TaMax can either visualize a single numerical attribute, a distribution, or multiple different attributes simultaneously.

We introduce two different approaches to visualize multiple attributes. In the **division approach**, each matrix cell is divided into n sub-cells and the attributes are visualized in them using a variety of different visual encodings, such as histograms, doughnut charts, or color coding. The number of sub-cells is automatically calculated based on the number of edge attributes that are visible at a given time. Fig. 2a shows an example where four different edge attributes are visualized in each cell.

The **overlay approach** shows two encodings on top of each other. A basic encoding can show arbitrary attributes and use different visual representations, a secondary encoding is achieved by overlaying as transparent rectangle, where the transparency is proportional to a numerical value, resulting in darker cells for higher values. Fig. 2b shows an example where histograms are overlaid with edge weights.

Both approaches have different pros and cons. When using subdividing, overall trends in the matrix can be difficult to spot due to clutter. When using overlays, the choice for base encodings are limited to those that do not interfere strongly with altered brightness, i.e., color-encoding is not suitable to be combined with overlays.

The analysis can choose which attributes are visible and how they are visualized. The adjacency matrix can be reordered based on the edge counts or other numerical edge attributes to reveal clusters and patterns in the graph.

TaMax supports data-driven aggregation of nodes. Rows and columns can be aggregated, expanded, collapsed, or sorted in the table view, the adjacency matrix is also rearranged to reflect the changes. Since the values in aggregated cells often are of vastly different scales than in individual cells, we introduce separate color-coding for each aggregation level, as shown in Fig. 1

Table View. In the table view, rows represent nodes and their attributes are visualized in the columns. The analyst has the ability to decide how to encode a given attribute. Currently, TaMax supports color saturation, area, histogram, pie chart, and text. Nodes can also be sorted based on the attributes.

TaMax allows the analyst to aggregate nodes based on any of the attributes in the table view. If the attribute is categorical, then the group is automatically generated. When the aggregation is done based on a numerical node attribute, we provide two options: either create two groups based on a threshold value, or aggregate into n groups. Both of these values can be specified by the user.

As the nodes are aggregated, we add new super-nodes to the graph. For these super-nodes, summary statistics based on the the attributes of their children are calculated and visualized in the table. The aggregation and addition of super-nodes creates a hierarchical structure which is visualized as a tree in the table. To view a high-level overview of the graph, the super-nodes can be collapsed individually. Figure 1 shows an example of migration between states where various regions are used to hierarchically aggregate the network. Multiple node attributes, such as the poverty level, average income, and the size of the population are shown in the table.

4 Future Work and Conclusion

We plan on revising and refining the types of supported encodings for the cells, and potentially add other attributes such as time series. We are also considering representing derived expressions in matrix cells, such as an aggregate of multiple values.

References

[1] B. Alper, B. Bach, N. Henry Riche, T. Isenberg, and J.-D. Fekete. Weighted graph comparison techniques for brain connectivity analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 483. ACM Press, Paris, France, 2013.

[2] N. Elmqvist, T. Do, H. Goodell, N. Henry, and J. Fekete. ZAME: Interactive Large-Scale Graph Visualization. In *2008 IEEE Pacific Visualization Symposium*, pp. 215–222, 2008. doi: 10.1109/PACIFICVIS.2008.4475479

[3] E. Kerzner, A. Lex, C. L. Sigulinsky, T. Urness, B. W. Jones, R. E. Marc, and M. Meyer. Graffinity: Visualizing Connectivity in Large Graphs. *Computer Graphics Forum*, 36(3):251–260, 2017.

[4] C. Nobre, M. Meyer, M. Streit, and A. Lex. The State of the Art in Visualizing Multivariate Networks. preprint, Open Science Framework, May 2019.