

MultiAggr: A Technique for Aggregating Multivariate Networks

Pranav Rajan*
University of Utah

Derya Akbaba†
University of Utah

Jack Wilburn‡
University of Utah

Roni Choudhury§
Kitware Inc.

Alexander Lex¶
University of Utah

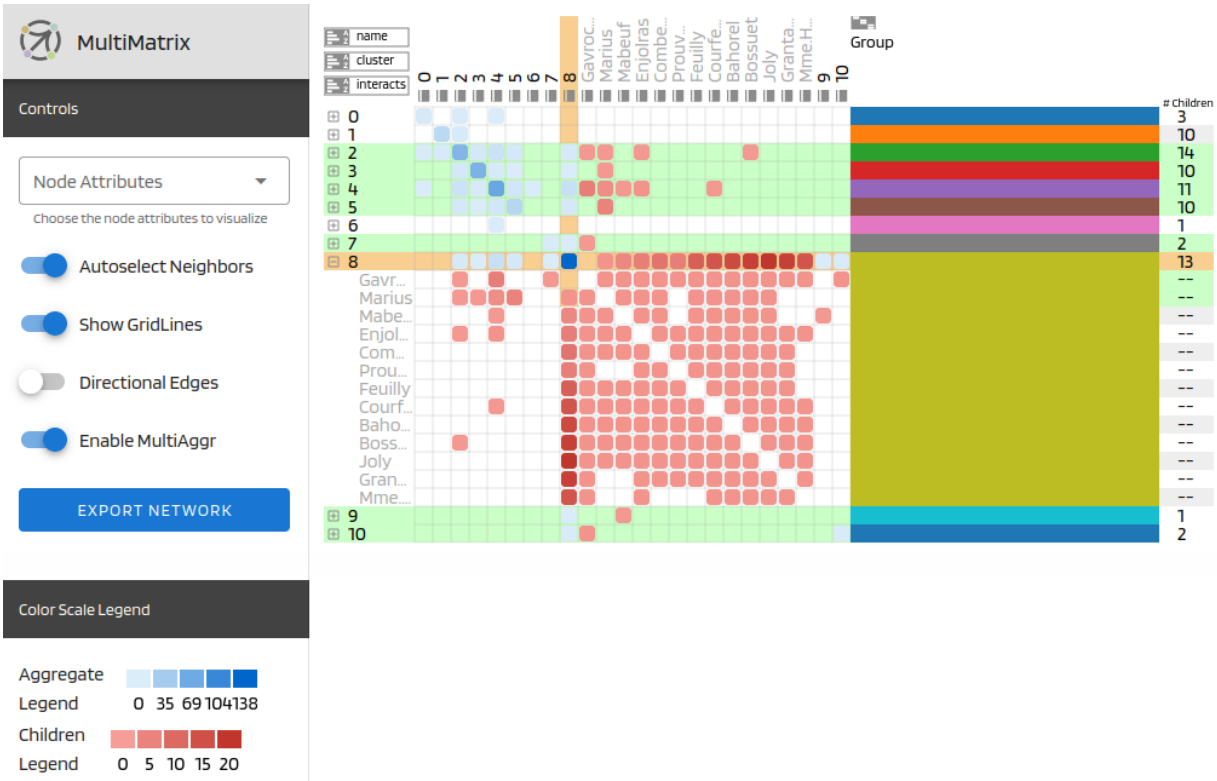


Figure 1: Aggregation of Les Misérables Character Relationship multivariate network using MultiAggr. The expanded AggrMatrix in this figure shows that the most character interactions were in Chapter 8 indicating a significant event in the novel. Upon expanding AggrMatrix for the Chapter 8 aggregated row, the expansion button reflects the act of expanding/retracting the row. Next to the right of the AggrMatrix is a count column keeping track of the number of children nodes that belong to an aggregated group. Dual color scale legends in the MultiMatrix Control Panel on the left of the AggrMatrix indicate the number of connections between non-aggregated nodes and aggregated nodes.

ABSTRACT

The analysis of dense multivariate networks (MVNs), including the data associated with the nodes and links, is an important task. Visualizing the attribute data in conjunction with the MVN data allows researchers to identify new questions and patterns about an MVN. Given limited screen space to visualize both MVN and attribute data concurrently, researchers have to perform data wrangling operations, including but not limited to aggregation, filtering, and sorting to subset a dense MVN for visual analysis. This work presents MultiAggr, a technique for aggregating MVNs across a single axis using categorical node attributes and the adjacency matrix layout of a network. MultiAggr can be used to obtain insight about high-level

overview patterns in different domains, including biology, social science, literature, and engineering.

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

1 INTRODUCTION

A popular layout for visual analysis of MVN data is the Node-Link Layout. However, the Node-Link Layout does not scale well and can lead to the visual clutter, and visual clutter as the number of nodes and links increase. Nobre et al. [4] identified the adjacency matrix layout as an effective layout for visualizing MVN data because nodes and attribute data can be represented using dedicated rows and columns as well as support multiple analysis tasks such as grouping and counting described by Lee et al. [3]. In practice, the adjacency matrix layout for visualizing MVN data requires quadratic screen space with respect to the number of nodes which limits the size of the MVN and associated edge attribute data that can be visualized without performing data preprocessing.

MultiAggr is inspired by Martin Wattenberg’s PivotGraph tool [6], Kerzner et al.’s Graffinity tool [2], and Nobre et al.’s MVN research [4] by focusing on visual aggregation for MVN data using

*e-mail: rajan.pranav320@gmail.com

†e-mail: derya@cs.utah.edu

‡e-mail: jwilburn@sci.utah.edu

§e-mail: roni.choudhury@kitware.com

¶e-mail: alex@sci.utah.edu

categorical node attributes and the adjacency matrix layout. While highly effective for the data domains they were developed for, PivotGraph and Graffinity are limited in the analysis features that allow users to develop new insights about their MVN data. To address these shortcomings, we developed MultiAggr, a domain agnostic visual aggregation technique for aggregating dense MVN data using categorical node attributes across a single axis using the adjacency matrix layout. MultiAggr is built using the latest web technologies, following a similar design to Graffinity but adds support for different types of MVN data, buttons that indicate to a user expand/retracting an aggregated row in a matrix, and a count column similar to the group count in the PivotGraph Control Panel.

2 FROM THEORY TO FEATURE

MultiAggr expands the capabilities of Graffinity and the MultiMatrix Application by adding detail-on-demand interactions, count summaries and dual visual encoding.

2.1 Detail-On-Demand Interaction: AggrMatrix

To show users an aggregated version of their MVN data using a single categorical node attribute, we developed AggrMatrix. AggrMatrix is an adjacency matrix that represents the result of performing MultiAggr on the MultiMatrix for a single categorical node attribute and gives users a high-level overview of their MVN data as show in Figure 2. Under the hood, AggrMatrix uses a rollup algorithm similar to PivotGraph’s rollup algorithm to aggregate MVN nodes into unique groups based up on the categorical node attribute selected by the user. Upon aggregation, the legend in the MultiMatrix Control Panel changes to a new color scale indicating to users that they are viewing aggregated data. Buttons next to the aggregated row labels enable users to perform details-on-demand interaction [5] through expanding and retracting aggregated groups of interest to see how groups of nodes are related to one another and other nodes in the multivariate network.

2.2 Data Summary: Count Column

We used Lee et al. and Nobre et al.’s research in data summary analytics tasks to develop the Count Column shown in Figures 1 and 2 [3, 4]. As part of our work to develop rollup for the adjacency matrix layout, we use *supernodes* [1] and *superlinks*, aggregated nodes and links for representing an aggregated MVN data set. Supernodes support computing derived grouped values to enable users to keep track of the number of items that belong to an aggregated group. Since it is not clear from the legend and the filled-in cell how many nodes belong to an aggregated group, we include the count column as a data summary.

2.3 Visual Encoding: Dual Color Encoding

Upon expanding the AggrMatrix, we use a dual-color encoding inspired by Graffinity to indicate aggregated and non-aggregated nodes. By dynamically expanding and retracting aggregated rows, users can identify relationships between aggregated and non-aggregated nodes and links to see how they relate to one another.

3 IMPLEMENTATION

MultiAggr is written in TypeScript, Vuetify, Vue, d3.js, and d3-legend as part of the MultiMatrix Application, a tool for performing interactive visual analysis of multivariate network data using an adjacency matrix layout as part of the MultiNet toolkit. Currently, MultiAggr supports small multivariate networks defined as less than one hundred nodes with categorical node attributes.

4 CONCLUSION AND FUTURE WORK

We wrote MultiAggr to demonstrate visual aggregation for categorical node attributes for domain agnostic multivariate network data from academic theory and prototypes to a production tool. We

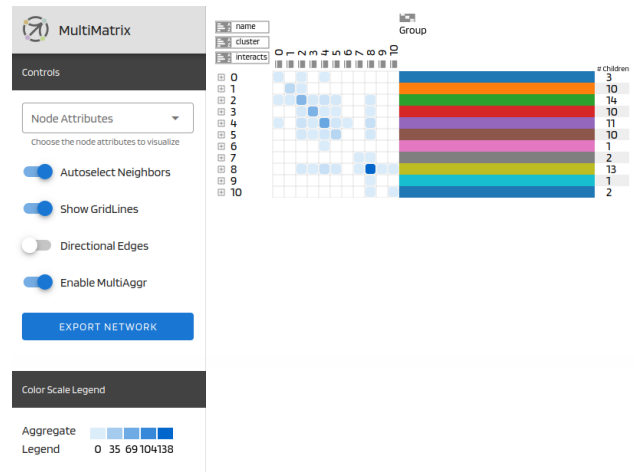


Figure 2: AggrMatrix for the Les Miserables Character Relationship multivariate network. The categorical attribute selected is the group attribute representing the different chapters of the novel which form the row and column labels. To the right of the AggrMatrix is the attribute column with different colored rectangles representing the unique groups.

hope that our work contributes to the advancement of multivariate network visualization and analysis and enables data analysts and researchers to discover new high-level patterns in their work in addition to making data wrangling and visualization more accessible. The MultiMatrix team plans to refine MultiAggr to support the tasks and goals of the MultiNet NSF Project and address the needs of the collaborators who are using MultiNet in their research.

ACKNOWLEDGMENTS

This work was supported by NSF OAC 1835904 and an REU supplement to the grant.

REFERENCES

- [1] A. Bigelow, C. Nobre, M. Meyer, and A. Lex. Origraph: Interactive network wrangling. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 81–92. arXiv:1812.06337, 2019. doi: 10.1109/VAST47406.2019.8986909
- [2] 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 (EuroVis)*, 36(3):251–260, 2017. doi: 10.1111/cgf.13184
- [3] 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*, pp. 1–5, 2006.
- [4] C. Nobre, M. Streit, M. Meyer, and A. Lex. The state of the art in visualizing multivariate networks. *Computer Graphics Forum (EuroVis)*, 38:807–832, 2019. doi: 10.1111/cgf.13728
- [5] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*, pp. 364–371. Elsevier, 2003.
- [6] M. Wattenberg. Visual exploration of multivariate graphs. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 811–819, 2006.