

Interactive Visualization of Time-Dependent Bipartite Graphs

Yahya Jabary*

TU Wien

ABSTRACT

This paper explores novel techniques for visualizing temporal bigraphs, with a specific focus on the “Media Transparency Database Austria”. Established in 2011 by Austrian law, this dataset records quarterly advertising expenditures of governmental organizations across various media outlets. Our study aims to develop an effective, interactive visualization that allows users to explore patterns, trends and anomalies.

We survey state-of-the-art approaches for visualizing dynamic graphs, including animated node-link diagrams, timeline-based representations and hybrid techniques. Building on these methods, we propose a novel interactive visualization that combines temporal bigraph representations. Our design aims to provide both clear overviews and detailed analysis capabilities, allowing users to explore the evolution of advertising expenditures over time and the relationships between public entities and media outlets. We discuss potential evaluation strategies for assessing the effectiveness of our visualization and outline expected outcomes. This work contributes to the field of dynamic graph visualization while demonstrating its practical application in government transparency and policy analysis.

Index Terms: Interactive Visualization, Temporal Graphs, Bipartite Graphs, Media Transparency, Public Policy, Government Spending

1 INTRODUCTION

Effective visualization of temporal graphs can reveal valuable insights such as patterns, trends, and anomalies that might otherwise remain hidden in raw data. This allows researchers, analysts, and decision-makers to quickly grasp evolving relationships, identify critical points of change, and make data-driven decisions.

In the context of global warming, multi-temporal scale visualizations of carbon emissions data can provide valuable insights for decision-makers, as demonstrated by Ma et al. [6]. In logistics, these visualizations can help optimize supply chain operations and identify inefficiencies [20] [14]. For social networks, they can illuminate user interactions and information spread. In public policy, temporal graph visualizations can aid in tracking disease spread, monitoring resource allocation and analyzing policy impacts [2]. These examples demonstrate how temporal graph visualizations compress temporal network data into intuitive representations. And as the volume and complexity of temporal data continue to grow across various domains, the development and application of effective visualization techniques for temporal graphs will likely become increasingly important.

As a practical example this paper explores various techniques for interactive visualization of time-dependent and bipartite graphs, using the “Media Transparency Database Austria” [5].

Time-dependent graphs or temporal graphs are structures that evolve over time, with nodes and edges changing as time progresses. These graphs are crucial for representing dynamic systems and processes that unfold chronologically. They allow us to observe patterns, trends, and changes in relationships between entities over different time periods [19]. One specific type of temporal graph that

has gained prominence is the bipartite graph. **Bipartite** graphs or bigraphs consist of two distinct sets of nodes, with connections only existing between nodes from different sets, but not within the same set. If the two subsets have equal cardinality (balanced bigraph) and each node in one subset is connected to every node in the other subset, the graph is called a complete bigraph [3]. These graphs are particularly useful for modeling relationships between two different types of entities, such as public authorities and their advertising expenditures. Due to their scale and complexity, time-dependent and bipartite graphs can be challenging to visualize effectively. To make sense of the “Media Transparency Database Austria”, which contains thousands of entries over multiple years, we need to design scalable and **interactive visualizations** that allow users to explore the data, for example, by filtering, zooming, and selecting specific time periods or entities.

The “Media Transparency Database Austria” [5] or **MT dataset** for short is a public repository that records the advertising expenses of public authorities in Austria. The MT dataset has gained popularity in Austria due to its role in increasing transparency around the use of taxpayers’ money for advertisements. Following allegations of misuse, the Austrian National Council passed the “Media Transparency Act” in 2011, requiring public institutions and certain companies to report their media expenditure quarterly. This legislation aims to make information easily accessible to citizens, thereby promoting open government initiatives and enhancing public accountability. It provides a rich dataset that can be represented as both a time-dependent and bipartite graph structure. By visualizing this data – although not the focus of this paper – we can gain insights into how public authorities allocate advertising budgets, which media outlets receive the most / least funding, how both the political and media landscapes evolve over time and mutually influence each other. This information can be used to evaluate the effectiveness of public advertising campaigns, identify potential biases and inform public policy decisions.

This paper aims to combine state-of-the-art techniques and approaches for designing interactive visualizations of time-dependent and bipartite graphs to propose a final design for the MT dataset. We will focus on clear, uncluttered overviews while maximizing information content and propose a design that allows users to interact with the data intuitively. Additionally, we will consider how our proposed visualization could be evaluated and what results we expect to achieve.

2 RELATED WORK

In this section, we will explore existing solutions for visualizing time-dependent and bipartite graphs, focusing on their strengths, weaknesses and suitability to the MT dataset. We can maximize suitability by focusing on existing research on our dataset, as they directly address our unique challenges. Fortunately the MT dataset has been a very popular choice for information visualization research in Austria starting from its inception in 2011. We will look at the relevant projects and papers in a chronological order (starting from the oldest) to understand the dialogue between them and how they have evolved over time.

*e-mail: yahya.jabary@tuwien.ac.at

2.1 Medientransparenz Projects (2013 - present)

In 2013, students at the Applied University of Joanneum were the first to work on the MT Dataset [7], which eventually evolved into the [Medientransparenz project](#) – the most comprehensive and up-to-date visualization of the MT Dataset. This project represents the cumulative effort of over 12 students’ course projects and master theses over the past 11 years, coordinated by Peter Salhofer. With generous funding from the Austrian government, it led to a direct collaboration with the Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR) and the Austrian Communications Authority (KommAustria), who are responsible for the MT Dataset [5].

On January 23 2024, this collaboration resulted in the project being forked into the [official Medientransparenz project](#), maintained by the RTR. The official Medientransparenz project is a restricted version of the original, featuring a different color palette and a limited set of features and datasets due to legal constraints applicable to KommAustria/RTR. For example, they are only permitted to visualize data from 2020 onwards. However, the original project remains unconstrained, actively maintained, and is self-described as a “Community Edition with extended data and functionality”, which will be the focus of our analysis and the one we will refer to as the Medientransparenz project from now on.

The Medientransparenz project is a single-page web-application based on Google’s material design framework. On the initial load of the webpage you have to wait an extended time period for it to load. It uses a vertical tab as a navigation menu, among others linking to the subpages “Overview”, “Top Values” and “Flows” which provide different views of the data.

The “Overview” page shows multiple views, accessible by scrolling. Each view provides access to 3 different visualizations of the same data subset by click on the corresponding icons: Line charts, bar charts and stacked bar charts. The goal of all these views is to provide a high-level overview of the data, showing the total spending of all entities over time, per type of spending. The default view is the bar chart displaying multiple color encoded bars per year, to enable the user to i.e. compare the advertising vs. funding related spendings. As you progress through the views, the subcategories of the data are further broken down, allowing for more detailed insights.

The “Top Values” page presents a complex hierarchical dataset using a modified pie chart structure, commonly referred to as a sunburst diagram or multilevel pie chart. A configuration tab on top of the page allows the user to select whether they’re interested in advertising spending or subsidies, funders or recipients, the cut-off rank and year as well as the different geographical regions. The design employs a radial layout to display hierarchical data, with the innermost ring representing the highest level of the hierarchy and subsequent outer rings showing more granular subcategories. However, it also demonstrates the challenges of presenting dense, multilevel information in a static, two-dimensional format which pose challenges in terms of readability and immediate comprehension. The use of color coding aids in category differentiation, but the large number of segments and varying text orientations may require significant cognitive effort from the viewer to fully interpret the data relationships, especially as the symbols overlap. Fortunately a tooltip and highlight feature is available to provide additional information on hover.

The third and final page, “Flows”, has a very similar filter configuration to the “Top Values” page, but instead of a sunburst diagram it uses a sankey diagram, a type of flow visualization that represents the distribution and flow of resources or information from one set of values to another. In this case the flow of money from funders to recipients, with the width of the lines representing the amount of money transferred. Additionally they’re great for temporal data, as they can show how the relationships between entities change over

time. The user can filter the data by year, region, and type of spending. This kind of visualization is excellent for bipartite graphs for obvious reasons, but it can be difficult to interpret for users unfamiliar with the concept – but it scales well and a tooltip is available to provide additional information on hover to help resolve visual ambiguities.

At this point it is also worth noting that although the sunburst diagram was preferred to the sankey diagram in the project maintained by the RTR and KommAustria due to its more intuitive nature, it is unable to capture the relationships in a bipartite graph as well as the sankey diagram can. The reason for this is unclear.

In conclusion this project provides a solid foundation for visualizing the MT dataset, with a focus on providing clear overviews and detailed insights. However, the lack of a direct comparison tool for actors and time spans, represents a missed opportunity for more sophisticated data analysis.

2.2 VISAR & CFAST Projects (2016)

In 2016, Alexander Rind and Wolfgang Aigner et al. collaborated on two highly relevant papers. The first paper was published in March 2016 with the VISAR team from the university of Stuttgart [1]. Later that year, in November, they teamed up again, this time including David Pfahler as one of the co-authors, to explore media transparency using multiple views [10], this time with the CFAST team. This collaboration sparked Pfahler’s interest in the subject, leading him to pursue it further in his master’s thesis. In 2019, Pfahler completed his thesis “Visualizing High-Dimensional Data with Hierarchically Aggregated Subsets” [8], supervised by Eduard Gröller, building upon the foundations laid in his earlier work with Rind and Aigner. Although not directly related to the MT dataset, Pfahler’s work is highly relevant as it explores the challenges of visualizing high-dimensional data, a common issue in the context of temporal and bipartite graphs.

The VISAR project “Visual Exploration of Media Transparency for Data Journalists: Problem Characterization and Abstraction” [1] describes the data structure and some basic statistical analysis, it does not propose or evaluate specific visualization methods. It doesn’t explore more advanced techniques that could handle the complexity of time-dependent bipartite graphs. Instead it suggests conducting interviews with data journalists to understand their needs and then categorizing tasks using established frameworks like Munzner’s what-why-how and the Data-User-Task Design Triangle. Although a step in the right direction, it isn’t relevant for our purposes.

The CFAST project “Exploring Media Transparency With Multiple Views” [10] implemented by Alexander Rind and David Pfahler [9] proposes a visualization using multiple coordinated views. The main techniques include bar charts for temporal and categorical aggregations, histograms for distributions, data tables with sparklines for details and a chord diagram for flows between entities. The chord diagram is an interesting alternative to the sankey diagram proposed in the Medientransparenz project, since they excel at displaying interconnections in a circular layout, making them ideal for showing bidirectional relationships and complex networks. Additionally they can also be combined with the previously suggested sunburst diagram, in a so called hybrid “sunburst-chord” diagram. Sankey diagrams on the other hand are better suited for visualizing directional flows and transitions between stages or categories, with the ability to clearly depict the magnitude of flow using weighted links.

The overall strengths of this paper include providing multi-granular views, interactivity and explorative analysis. However by providing 7 dense visualizations on a single page, the authors risk overwhelming the user with information. The chord diagram in particular can be difficult to interpret for users unfamiliar with the

concept. A single “hide” button for all visualizations followed by a soft animation is a missed opportunity to reduce the sensory overload.

2.3 VDA Student Projects (2017)

Jüttner et al. [4] present an interactive visualization approach specifically designed for the MT dataset that directly addresses our unique challenges. They describe a visualization approach in which different aspects of the data are simultaneously presented through interactive and interlinked sub-visualizations, called views. These multiple views work in concert, all showing different aspects of the same data and being updated in sync when users interact with any of them. This is commonly referred to as a “multiple coordinated views” or “*coordinated multiple views*” (CMV) [11] visualization and aligns well with the popular design principle of “Eyes beat memory” by Tamara Munzner. This principle suggests that users can more effectively understand data when they can see it, rather than having to remember it from a previous view. Their visualization concept provides both broad overviews (i.e. overall spending patterns) and detailed insights (i.e. drill down into specific actors or time periods). It consists of elements, including an overview bar chart, an alphabetically ordered, searchable list of actors, a tree-like graph displaying connections between selected actors and detailed bar charts showing quarterly spending and media-specific expenses.

This ability to interactively filter data, search for specific actors, and update charts based on user selections is a great strength of their approach. It allows users to explore the data in a non-linear fashion, focusing on specific aspects of the data that interest them.

However, the approach also has some limitations. The authors acknowledge performance issues especially on startup as their web-based `d3.js` implementation doesn’t scale well with the given dataset size. This could potentially hinder user experience, especially for users with slower internet connections or less powerful devices. Additionally, the lack of advanced features like brushing, zooming, and direct comparison tools for actors and time spans, as mentioned by the authors, represents a missed opportunity for more sophisticated data analysis.

In conclusion while Jüttner et al.’s approach provides a solid foundation through its coordinated multiple views design, it doesn’t leverage the temporal and bipartite nature of the data to its full potential.

After some research, we found that Jüttner et al.’s work was part of a student project at the University of Vienna in 2016 WS [16] before being published in 2017. This context is important to note as it suggests that other students might have also worked on similar projects during that time. By modifying the URL suffix of the original project, we were able to find several other student projects from the same year that also explored the MT dataset and used interactive CMV visualizations that allow to drill down into the data, but each with unique approaches and interesting insights. We also learned that the deadlines for these projects were on the 21st of January 2017, which is why they were all published around the same time. Two of these projects, in particular, caught our attention.

One project [17] used a bundled edge graph to provide a detailed view of individual transactions between entities, a packed bubble chart and treemap to provide overviews of the highest transactions and relative proportions and a line chart that allows users to see trends over time. Additionally each graph would highlight connections on mouseover in the bundled edge graph.

Notable weaknesses however include the lack of cross-filtering between the bundled edge graph and the other visualizations, requiring users to apply filters separately. This can be cumbersome and potentially confusing. The default view can also be overwhelming due to the large amount of data, making it initially difficult for users to gain clear insights. Performance issues with loading large datasets using `d3.js`, particularly in the bundled edge graph, are

another weakness. Also the decision to display nodes as circles without text descriptions when there are more than 250 nodes is a compromise that may not be ideal for all use cases.

Another project [18] took a different approach, aiming for simplicity and information density. They used a matrix view to show the relationships between entities and media outlets. The project also included a stacked bar chart to show the distribution of spending across media outlets and a timeline view to show spending trends over time. The matrix view in particular is a unique, effective and compact way to visualize bipartite relationships, providing a clear and information-dense representation of the data.

The reliance on basic chart types however limits the ability to show more complex relationships or patterns in the data which may be appropriate for a general audience but could limit its utility for expert users. The stacked bar charts can be difficult to interpret precisely. The matrix view, while information-dense, may be overwhelming for some users.

2.4 Dynamic BiCFlows (2018 - 2019)

In 2018, Steinböck et al. [13] laid the foundation for the “Dynamic BiCFlows” project, which was improved upon in 2019 by Wolfgang Knecht and Johannes Sorger from the Complexity Science Hub Vienna [12] and launched in 2020. All of these projects were co-supervised by Manuela Waldner from TU Wien, with her most recent thesis supervision on bipartite graphs occurring in 2020 on “Interactive exploration of large time-dependent bipartite graphs [...]” [15]. Her research group has been at the forefront of bipartite graph visualization in Austria.

The first mentioned paper by Steinböck et al. from 2018 [13] proposes a technique called “BiCFlows” that combines hierarchical aggregation through biclustering with filtering in adjacent lists. The key idea is to use biclustering to group nodes into meaningful clusters, which are then visualized as adjacent vertical lists. Users can interactively drill down into clusters to reveal more details. The hierarchical aggregation helps reveal patterns and relationships that may not be apparent when looking at individual nodes. The authors found that users spent more time exploring the data and discovered more unexpected insights compared to a baseline filtering-only approach. However users also perceived it as more complex than the simpler filtering interface, at least initially. The evaluation showed that users typically only drilled down one or two hierarchy levels, potentially missing some of the lower-ranked nodes and relationships. The clustering approach may also not always produce meaningful groupings, especially if there is low modularity in the data. Additionally, the biclustering algorithm can be computationally expensive for very large graphs, potentially limiting scalability and also does not have explicit support for temporal aspects. For time-dependent bipartite graphs, additional visual encodings or interaction techniques may be needed to convey temporal patterns and changes over time.

The paper by Knecht and Sorger from 2019 [12] isn’t optimized for the MT dataset. It proposes an immersive virtual reality (VR) approach for visualizing and exploring large dynamic networks. The key visualization technique suggested is a dual-perspective system that combines an exocentric overview of the entire network with an egocentric detail view that allows users to immerse themselves within the network structure. This first-person view enables detailed inspection of dense areas and discovery of connections that may be obscured in the overview. The paper also proposes temporal navigation controls to explore the network’s evolution over time. For the overview, the paper suggests rendering the network as a 3D node-link diagram using a force-directed layout. Nodes are represented as spheres of varying sizes, while edges are shown as tubes, with arrows indicating directionality. This provides a familiar representation that expert users can easily relate to conventional 2D

network visualizations. For the MT dataset the overview would allow analysts to quickly identify key players and overall spending patterns. The detail view could be particularly useful for exploring complex relationships between government entities and media companies that may be difficult to discern in a 2D layout. The temporal navigation would enable tracking of spending trends over time. However, the VR approach may not be suitable for all users, especially those without access to VR hardware or who are not comfortable with immersive environments. The dual-perspective system may also require additional cognitive effort to switch between views, potentially slowing down the analysis process.

The second mentioned paper by Steinböck et al. from 2020 [19]

This paper proposes a novel visualization technique called Dynamic Bipartite Cluster Flows (Dynamic BiCFlows) for exploring large time-dependent bipartite graphs. The technique combines hierarchical aggregation and filtering to visualize the data as linked lists with nested time series.

For the Media Transparency Database Austria dataset, Dynamic BiCFlows offers several strengths. It allows users to gradually drill down from an overview to detailed levels, revealing connections between legal entities and media organizations at different granularities. The hierarchical aggregation helps manage the scale of the dataset, which contains thousands of nodes and edges. The nested time series visualizations enable users to observe temporal patterns in advertising expenditures over different time periods. The interactive exploration capabilities, including drill-down, filtering, and details-on-demand, encourage users to spend more time exploring the data and discover unexpected insights.

However, the technique also has some limitations for this dataset. The user study revealed that participants found BiCFlows more complex to use initially compared to a simple filtering approach. The hierarchical aggregation may obscure some smaller entities or transactions that users might be interested in finding directly. Additionally, the biclustering approach used for aggregation may not always produce meaningful groupings that align with users' mental models of the data.

In terms of general applicability to time-dependent bipartite graphs, Dynamic BiCFlows shows promise. The technique is designed to scale to thousands of nodes and edges, making it suitable for large datasets. The combination of biclustering and time series clustering provides flexibility in exploring both structural and temporal aspects of bipartite graphs. The linked list visualization with nested time series offers an intuitive way to represent bipartite relationships and their evolution over time.

However, the effectiveness of the technique may depend on the specific characteristics of the dataset. Graphs with low modularity may result in less meaningful clusters and more visual clutter. The time series clustering approach assumes that temporal patterns are an important aspect of the data, which may not always be the case for all time-dependent bipartite graphs.

Overall, Dynamic BiCFlows presents an innovative approach to visualizing large time-dependent bipartite graphs, offering a balance between overview and detail. While it shows particular strengths for datasets like the Media Transparency Database Austria, its general applicability to other time-dependent bipartite graphs warrants further investigation across diverse domains and data characteristics.

The vertical length of the flow nodes distorts the timeseries, making it difficult to compare the flow values across different time periods.

3 VISUAL ENCODING AND INTERACTION DESIGN

Next we will build upon the existing research to propose a novel visualization technique for the MT dataset.

<https://dl.acm.org/doi/abs/10.1145/3544548.3581119>

"Sunburst-Chord" or "Hierarchical Edge Bundling" diagram

Horizon Graphs

4 IMPLEMENTATION

5 EVALUATION

REFERENCES

- [1] W. Aigner, C. Niederer, M. Sedlmair, A. Rind, R. Gutounig, J. Ausserhofer, et al. Visual exploration of media transparency for data journalists: Problem characterization and abstraction. 2016. <https://visvar.github.io/pub/niederer2016ffh.html> and https://research.fhstp.ac.at/content/download/89494/file/2016_FFH_VALiD_Niederer.pdf. 2
- [2] W. Chung and V. S. Lai. A temporal graph framework for intelligence extraction in social media networks. *Information & Management*, 60(4):103773, 2023. 1
- [3] R. Diestel. *Graph theory: Springer graduate text gtm 173*, vol. 173. Reinhard Diestel, 2012. 1
- [4] R. Jüttner, B. Potuzak, and R. Miftari. Media transparency-an interactive visualization of advertisement money flows between governmental organizations and austrian media companies. 2017. 3
- [5] R. . KommAustria. Media Transparency Database Austria. https://www.rtr.at/medien/was_wir_tun/medientransparenz/VeroeffentlichteDaten/veroeffentlichungen.de.html, 2024. [Online; accessed 01-July-2024]. 1, 2
- [6] M. Ma, P. Xie, F. Teng, B. Wang, S. Ji, J. Zhang, and T. Li. Hist-gnn: Hierarchical spatio-temporal graph neural network for weather forecasting. *Information Sciences*, 648:119580, 2023. 1
- [7] F. J. G. mbH. mehr! medien transparenz. <https://www.medien-transparenz.at/impress>, 2011. [Online; accessed 01-July-2024]. 2
- [8] D. Pfahler. Visual summaries of high-dimensional data using hierarchical aggregated data-subsets. 2019. [Online; accessed 01-July-2024]. 2
- [9] A. Rind and D. Pfahler. MTDB2. url-<https://github.com/VALIDproject/mtdb2>, 2016. [Online; accessed 01-July-2024]. 2
- [10] A. Rind, D. Pfahler, C. Niederer, and W. Aigner. Exploring media transparency with multiple views. In *FMT*, pp. 65–73, 2016. <https://repositum.tuwien.at/handle/20.500.12708/56612> and <https://www.cvast.tuwien.ac.at/topics/media-transparency-dashboard> and <https://github.com/VALIDproject/mtdb2>. 2
- [11] J. C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *Fifth international conference on coordinated and multiple views in exploratory visualization (CMV 2007)*, pp. 61–71. IEEE, 2007. 3
- [12] J. Sorger, M. Waldner, W. Knecht, and A. Arleo. Immersive analytics of large dynamic networks via overview and detail navigation. In *2019 IEEE international conference on artificial intelligence and virtual reality (AIVR)*, pp. 144–1447. IEEE, 2019. <https://users.cg.tuwien.ac.at/~waldner/bicflows/> and https://github.com/complexity-science-hub/dynamic_bicflows and <https://repositum.tuwien.at/handle/20.500.12708/57887.3>
- [13] D. Steinböck, E. Gröller, and M. Waldner. Casual visual exploration of large bipartite graphs using hierarchical aggregation and filtering. *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pp. 1–10, 2018. <https://repositum.tuwien.at/handle/20.500.12708/57540.3>
- [14] R. Tamilmani and E. Stefanakis. Modelling and analysis of semantically enriched simplified trajectories using graph databases. *Advances in Cartography and GIScience of the ICA*, 2019. 1
- [15] K. Unger. *Interactive visual exploration of large bipartite graphs using firework plots*. PhD thesis, Wien, 2020. 3
- [16] Visualization and V. U. o. T. Data Analysis group (VDA). Student Reports on Visualization Projects. http://vda.univie.ac.at/Teaching/Vis/16w/Final_Report/10.pdf, 2017. [Online; accessed 01-July-2024]. 3
- [17] Visualization and V. U. o. T. Data Analysis group (VDA). Student Reports on Visualization Projects. <http://vda.univie.ac.at/>

- [Teaching/Vis/16w/Final_Report/03.pdf](#), 2017. [Online; accessed 01-July-2024]. 3
- [18] Visualization and V. U. o. T. Data Analysis group (VDA). Student Reports on Visualization Projects. http://vda.univie.ac.at/Teaching/Vis/16w/Final_Report/07.pdf, 2017. [Online; accessed 01-July-2024]. 3
- [19] M. Waldner, D. Steinböck, and E. Gröller. Interactive exploration of large time-dependent bipartite graphs. *J. Comput. Lang.*, 57:100959, 2020. <https://repositum.tuwien.at/handle/20.500.12708/141146> and <https://www.cg.tuwien.ac.at/research/publications/2020/waldner-2020-tbg/>. 1, 4
- [20] H. Yang, F. Zhou, G. Lin, M. Lin, and S. Lin. Animated mesh simplification based on motion features in visual sensor networks. *International Journal of Distributed Sensor Networks*, 15, 2019. 1