

Designing Scalable Storylines for Hypergraphs Exploration

Master 2 Internship

Anastasia Bezerianos and Vanessa Peña-Araya
{anab@lri.fr, vanessa.pena-araya@inria.fr}

Internship period: starting in March/April 2023, 5 or 6 months duration.

Supervisors: Anastasia Bezerianos and Vanessa Peña-Araya.

Location: ILDA, Université Paris-Saclay, building 660.

In our past contact with data journalists we've seen that they are interested in analyzing the relationships among politicians, whether they visit particular locations or if they are linked to certain organizations (e.g. political parties, enterprises, etc). These relationships help them understand recent news events, like why a politician won an election or why they quit their political party.

Visualizing these datasets is an emerging research field [2]. Our HyperStorylines [3] tool is one of the most recently published visualization techniques to explore them. It is based on Storylines visualizations where people are represented by lines that evolve over the horizontal axis that represents time. HyperStorylines generalizes Storylines visualizations by allowing users to create custom views and see the relationships of any two types of entities, instead of just people over time. Additionally, a third type of entity can be visible by using interaction. Figure 1 shows some examples of custom views of HyperStorylines.

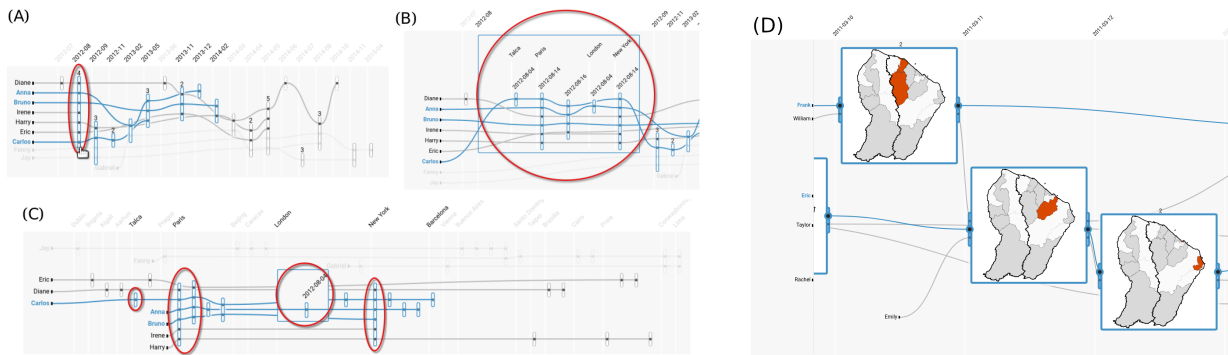


Figure 1: (A-C) HyperStorylines visualization with three views of a dataset. (A) Shows a view of people (represented by lines) that evolve along the horizontal axis, that here represents time (aggregated by months). Small vertical bars are relationships, positioned in the intersection of both axes of the entities that compose them. Relationships can have zero or more internal nested entities (a third type of entity), which can be seen by interactively expanding them (B). (C) Shows the stories of people related by locations instead of time (time is the nested entity). The red circles across images indicate where the entities that contribute to the highlighted relationship in (A) appear in the other views. (D) Shows an example of GeoStorylines, a visualization that shows the geographical context of these relationships.

Internship goal:

The volume of real-world datasets makes the resulting views complex and hard to analyze. The goal of this internship is to design scalable Storylines visualizations that summarized views and allow users to observe high level patterns. Based on our conversations with journalists we are interested in seeing recurrent patterns in space and time as well as relationships that include many people.

More context:

As inspiration, there are some visualizations in the literature that address similar problems. For example, Reda *et al.* [5] visualizes the evolution of communities structures over time (Figure 2 top). On the other hand, Pister *et al.* [4] (Figure 2 bottom) show the results of different clustering algorithms to see the relationships among entities over time. However, most of the available visualizations mainly focus on identifying and following groups of people that evolve

over time and do not consider the types of patterns we are interested in such as recurrent events in space and time or events that involve multiple people.

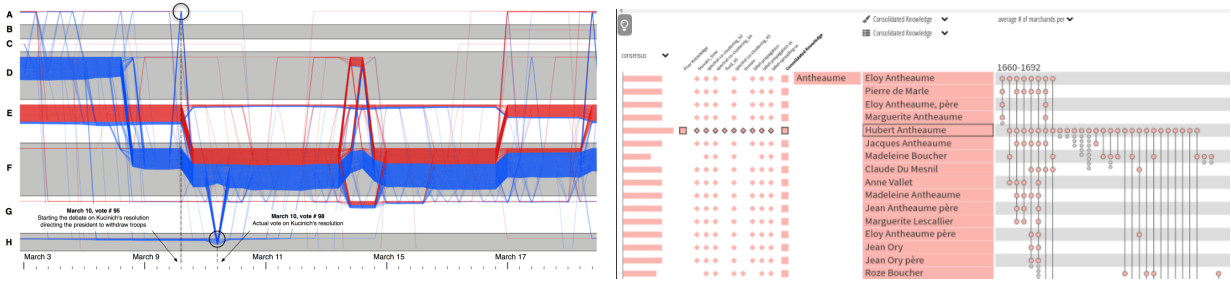


Figure 2: Two visualizations that show groups of entities evolving over time. On the top is the work of Reda *et al.* [5], a visualization that shows the evolution of community structures in dynamic social networks. On the bottom, we see PK-clustering [4] a visualization that shows the results of different clustering algorithms to see groups of entities over time.

Work plan:

The work of the internship will be divided in three main tasks:

1. Make a literature review about visualization of groups of entities evolving over time.
2. Conduct design workshops to generate alternatives for scalable Storylines is
3. Creation of a design space (taxonomy) of possible designs and tasks they support.
4. **IF** time permits, implementation + evaluation for most promising designs

Requirements for Applicants: Knowledge of user evaluation and prototyping methods. Any past experience in web development and information visualization is a big plus.

References

- [1] AGARWAL, S., AND BECK, F. Set streams: Visual exploration of dynamic overlapping sets. *Computer Graphics Forum* 39, 3 (2020), 383–391.
- [2] FISCHER, M. T., FRINGS, A., KEIM, D. A., AND SEEBACHER, D. Towards a survey on static and dynamic hypergraph visualizations, 2021.
- [3] PEÑA-ARAYA, V., XUE, T., PIETRIGA, E., AMSALEG, L., AND BEZERIANOS, A. Hyperstorylines: Interactively untangling dynamic hypergraphs. *Information Visualization* 0, 0 (0), 14738716211045007.
- [4] PISTER, A., BUONO, P., FEKETE, J.-D., PLAISANT, C., AND VALDIVIA, P. Integrating prior knowledge in mixed-initiative social network clustering. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1775–1785.
- [5] REDA, K., TANTIPATHANANANDH, C., JOHNSON, A., LEIGH, J., AND BERGER-WOLF, T. Visualizing the evolution of community structures in dynamic social networks. *Computer Graphics Forum* 30, 3 (2011), 1061–1070.