Information Visualization Assignment

MetroSets: Visualizing Sets as Metro Maps

Ben Jacobsen, Markus Wallinger, Stephen Kobourov, and Martin Nollenburg

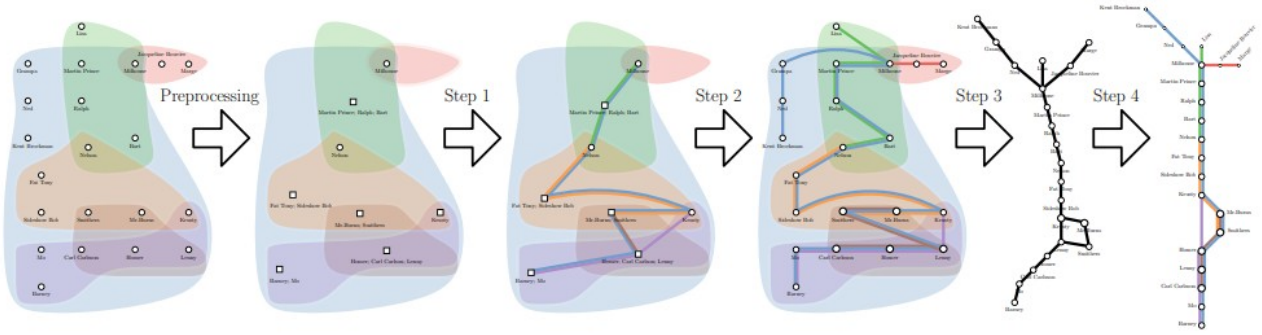

Fig. 1: The MetroSets pipeline. The input set system (here: characters of *The Simpsons*) is first compressed into a combinatorially equivalent smaller instance; Step 1 creates an optimized path support graph; Step 2 reinserts temporarily discarded elements; Step 3 creates an initial layout of the support graph; finally Step 4 schematizes the layout as a metro map and places the labels.

The authors describe MetroSets as a flexible and modular system for visualizing set systems using the schematic metro map metaphor. The sets are represented by metro lines and set elements are represented by metro stations, with elements that belong to multiple sets corresponding to interchange stations; see Figure 1. This approach leverages the familiarity of metro maps and scales well for medium-sized datasets with over a dozen sets and hundreds of elements. Set systems are used in many applications. Such analysis can be aided greatly by effective data visualization. The work focuses on datasets which are too large to be effectively visualized with Euler Diagrams, but not so large that they require a full-fledged visual analytics system.

The paper's findings are followed: 1. The running time of MetroSets is most strongly influenced by the number of vertices. 2. For most metrics used in experiments, the number of hyperedges is the most important factor. 3. For most metrics, the choice of support or insertion does not have a large effect on the final score, and the impact of the schematization algorithm is limited to octolinearity and edge uniformity.

The paper's limitations are followed: 1. Due to its quadratic runtime, in the online setting the system is practically limited to inputs with no more than 200 vertices and 20 hyperedges. 2. There is a lack of control over aspect ratio of the drawing area, poor outputs when any vertex must have degree grater than 8, and a limit of 20 on the number of hyperedges before line colors must be reused. 3. Space utilization is also not optimized, although adding a 'compaction' step to the pipeline might be possible.

In my personal opinion, the strength of the paper is that the topic is simple but novel. The visualization method to visualize metro is very familiar to the audience but the method has been rarely used except describing it. However, the authors adopt the style to represent other data in many different fields. However, there are several critical weaknesses that the topic

intrinsically reveals. As aforementioned, the method can never be free from spatial confinement. This is also an issue for many other metro maps and this graph-like feature which totally relies on colors and vertices. For the problem, I suggest attaching animations or motions such as zooming or brushing to focus on certain point of the whole graph in a limited spatial or color option and overcome the limitations. Plus, another big issue of the paper is that the authors did not ran human subjects study evaluating MetroSets in comparison to alternative set visualization systems. The paper only asserts that the method is the most suitable method for data of moderate size but there is no valid proof for it. If the authors expend the topic by covering the weaknesses revealed in the paper, the metro-like visualizing method could be widely adopted in many cases.