10

Scalability Considerations for Multivariate Graph Visualization

T.J. Jankun-Kelly, Tim Dwyer, Danny Holten, Christophe Hurter,
Martin Nöllenburg, Chris Weaver, and Kai Xu

Scalability in visualization is a challenge: How do we choose to show more items than can be easily rendered upon a screen or understood by a human effectively? Multivariate graph visualization adds additional wrinkles in that nodes and edges are no longer atomic entities. Rather, they are repositories for further rich information. In information seeking, the mantra attributed to Ben Shneiderman succinctly outlines a path to visual scalability: "Overview first, zoom, then details-on-demand" [69]. While this is good guidance, naively presenting the whole universe of data as an initial "overview", leads to dense, unreadable displays (Fig. 10.1). To provide insightful visualizations at large scale for multivariate graphs, we must understand what our visual, cognitive, and architectural limits are, then explore approaches to mitigate these limitations. Detailed views must offer useful affordances for navigation to other views. The goals of this chapter are to identify the challenges and the state-of-the-art in these areas.

At large scale, dense multivariate graphs devolve into *hairballs* (dense collections of nodes with heavily over plotted edges) or *snowy wastes* (highly populated matrix diagrams with visually random structure) if the entire structure is shown (Fig. 10.1). Perceptual and cognitive psychology outline what human visual and mental limitations interfere with understanding such dense views; additionally, there are hardware factors which band the amount of graph data that can be rendered in a timely matter. By understanding these limitations, outlined in Sect. 10.1, designers can utilize the strategies explored in Sect. 10.2 to show only what is needed when it is needed. Use of these strategies, and further studies on the limits of scalability, are also presented in Sect. 10.3 as a means to guide further research. We conclude with a summary of challenges in scalable, multivariate graph visualization.