9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Real-time visualization of large networks and embeddings

Keywords: network visualization, embedding visualization, exploration tool, science of science, social media

Extended Abstract

The field of computational social science is heavily reliant on the study of complex systems and the application of network science. In their various forms, networks are the fundamental building blocks for representing and analyzing the interconnections within complex systems. Such systems, be it social media interactions, scientific collaborations, or the diffusion of innovations, often exhibit complex topologies consisting of many nodes and edges, along with metadata and other features. Interactive network visualization plays a crucial role in understanding these systems and uncovering hidden patterns. It provides a holistic perspective and enables intuitive interpretations of underlying phenomena. Network visualization can also be applied to understand processes, such as machine learning pipelines or simulations, as well as to disseminate results to a wider audience [1, 2, 3]. However, current tools often fall short in handling large networks with over 10,000 nodes in real-time due to limited rendering capabilities and the absence of continuous layout algorithms [1].

To tackle the challenge of visualizing intricate and large networks, we introduce a new network exploration tool capable of visualizing datasets with millions of nodes through GPU-based rendering and continuous force-directed layouts. This is accomplished by reducing the number of calls and information transmitted from the CPU to the GPU and implementing instantiated geometry [4]. Our tool employs billboard rendering [4], which utilizes a single geometry source to render all the glyphs and lines on the screen. Each instance is then aligned to face the camera through a vertex shader on the GPU (as shown in Figure 1a). Nodes and edges are rendered using signed distance functions (SDF) [5], which enables the use of various shapes and effects for the glyphs (as illustrated in Figure 1b). This technique allows for the creation of pixel-perfect 2D and 3D shapes. The shapes, colors, and other visual characteristics of the glyphs are drawn on top of the quads using a fragment shader. Since the fragment shader is executed for each pixel on the screen, there is no degradation in quality when rendering complex shapes. Moreover, texture atlases can be utilized to display images or dynamic textures for nodes or edges.

Selecting nodes and edges in large networks can be challenging with traditional geometric methods (such as ray casting) as they perform poorly with increasing size and complexity of the geometry being rendered. To overcome this limitation, our tool integrates a high-performance picking system that works seamlessly with the GPU-based rendering pipeline [6]. This technique is based on rendering the scene to an external frame buffer with minor modifications to the fragment shader. Instead of drawing colors for display, the pixel colors encode the index of the nodes or edges. Picking is carried out by retrieving the rendered pixel in the picking frame buffer at the position of a cursor click or hover event and decoding the index from the color. This hardware-based picking allows real-time frame-by-frame interactions with nodes or edges, such as hovering, dragging, and filtering (as shown in Figure 1d).

The GPU-based pipeline of the tool enables the implementation of advanced effects to enhance insights and interactions with data. For instance, kernel density estimation in real-time allows exploring network properties or attributes assigned to nodes (as shown in Figure 1d). To demonstrate the effectiveness of this tool, we applied it to explore scholarly networks and embeddings derived from large portions of the OpenAlex dataset (as shown in Figure 1f). Also, the tool was integrated into a visualization pipeline to analyze large social media networks on the web, specifically showcasing user interactions and hash term co-occurrence networks from Twitter datasets and API (Figure 1f)).

The tool includes an API and interactive features, allowing users to search, filter, and highlight nodes or edges based on their attributes. Our tool outperforms existing open-source solutions for network visualization, including Graphviz, Gephi, Cytoscape, graph-tool, igraph, networkx, Graphia, and 3d-force-graph, being capable of rendering networks with more than 50,000 nodes on modest hardware and more than one million nodes on better hardware. In addition to networks, projections of embeddings can also be explored with the tool, which also provides capabilities to display the neighbors of entities in the original space. In conclusion, our framework provides means to handle large networks with millions of nodes and facilitates real-time visualization and exploration of dynamic complex networks. However, its main objective is not to replace existing open-source solutions but to provide an alternative backend that can be used standalone or integrated into existing or new software and platforms.

References

[1] J. Park, I. B. Wood, E. Jing, A. Nematzadeh, S. Ghosh, M. D. Conover, and Y.-Y. Ahn, "Global labor flow network reveals the hierarchical organization and dynamics of geo-industrial clusters," *Nature communications*, vol. 10, no. 1, p. 3449, 2019.

[2] T. Nishikawa and A. E. Motter, "Discovering network structure beyond communities," *Scientific reports*, vol. 1, no. 1, pp. 1–7, 2011.

[3] B. Schäfer, D. Witthaut, M. Timme, and V. Latora, "Dynamically induced cascading failures in power grids," *Nature communications*, vol. 9, no. 1, pp. 1–13, 2018.

[4] L. Wagner, D. Limberger, W. Scheibel, M. Trapp, and J. Döllner, "A framework for interactive exploration of clusters in massive data using 3d scatter plots and webgl," in *The 25th International Conference on 3D Web Technology*, pp. 1–2, 2020.

[5] C. Green, "Improved alpha-tested magnification for vector textures and special effects," in *ACM SIGGRAPH 2007 courses*, pp. 9–18, 2007.

[6] M. Tomas, E. Haines, N. Hoffman, A. Pesce, M. Iwanicki, and S. Hillaire, *Real-time rendering*. CRC Press, 4 ed., 2018.

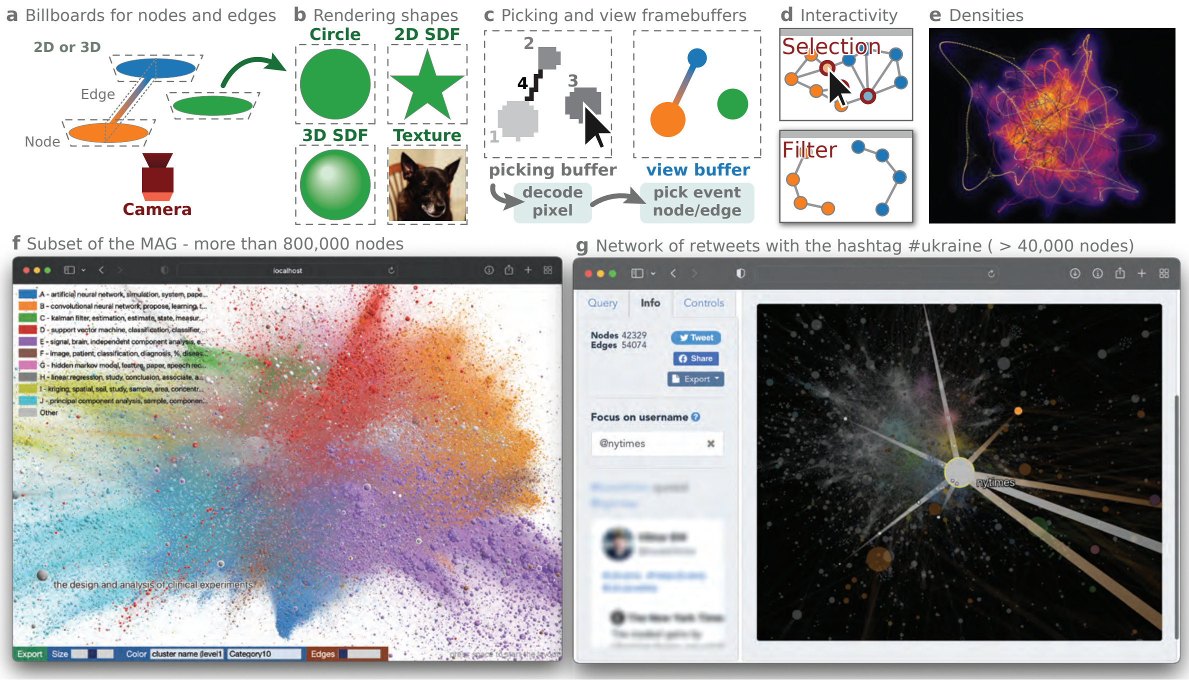

Figure 1: (a) GPU-based rendering via billboards (simple geometry always facing the camera). (b) Markers for nodes rendered using signed distance fields or texture atlases. (c) picking nodes and edges is accomplished in the GPU using a picking framebuffer. (d) Example of selection and filter interactions. (e) Kernel-based Density of nodes, edges or any attribute can be rendered in real-time. (f) Example of a 3D projection of a node2vec embedding of papers in the Artificial Intelligence field, obtained from the OpenAlex dataset. (g) Screen capture of the developed network tool used to explore large social media networks.