

Escalabilidad visual en grandes grafos*

Dana K. Urribarri
dku@cs.uns.edu.ar

Silvia M. Castro
smc@cs.uns.edu.ar

Sergio R. Martig
srm@cs.uns.edu.ar

Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica
Departamento de Ciencias e Ingeniería de la Computación
Tel: (0291) 459-5135 Fax: (0291) 459-5136
Universidad Nacional del Sur
Bahía Blanca, Buenos Aires, Argentina

Resumen

El crecimiento incesante del volumen de información y la necesidad de analizar e interpretar los datos para extraer conocimiento útil de ellos, hacen que la escalabilidad, en general, y la escalabilidad visual, en particular, sean de gran importancia al momento de evaluar y diseñar técnicas de visualización.

El principal objetivo del trabajo de investigación que está en desarrollo la exploración de nuevas técnicas que incrementen la escalabilidad visual de las representaciones existentes en el campo de la visualización de grafos, así como las interacciones que la soportan.

Palabras Claves: escalabilidad visual, interacciones, visualización de grafos, visualización de información.

1. Introducción

La información crece constantemente, no sólo en volumen sino también en diversidad. Es común encontrar información textual, por ejemplo noticias, e-mails y páginas webs, bases de datos, imágenes, como imágenes satelitales, datos relevados de sensores y videos, entre otros. Lamentablemente, estos datos no sólo son diversos, voluminosos y dinámicos sino que también son ambiguos e inciertos. Estas características hacen que se vuelva un desafío la tarea de proveer mecanismos para visualizar relaciones en la información y permitir a los analistas obtener conocimiento nuevo de los datos.

A pesar de la tecnología actual, las herramientas existentes no asisten lo suficiente a las habilidades humanas (que no cambian significativamente a lo largo del tiempo) para el procesamiento de volúmenes de información en constante aumento. Esto obliga a prestar mayor importancia a la *escalabilidad* y sus diferentes facetas [6]:

- *Escalabilidad visual*: capacidad de las representaciones visuales y de las herramientas de visualización de mostrar efectivamente grandes conjuntos de datos.
- *Escalabilidad de información*: capacidad de extraer información relevante de grandes flujos de datos.

*El presente trabajo fue parcialmente financiado por PGI 24/N015 y PGI 24/ZN12, Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur, Bahía Blanca, Argentina.

- *Escalabilidad de software*: capacidad de un sistema de software de manejar interactivamente grandes conjuntos de datos.
- *Escalabilidad analítica*: capacidad de los algoritmos matemáticos de manejar en forma eficiente grandes conjuntos de datos.

Sin olvidarnos de la escalabilidad de software, de gran importancia al momento de implementar herramientas de visualización, nosotros nos abocaremos a trabajar esencialmente en escalabilidad visual.

2. Escalabilidad visual

Dadas las limitaciones de la tecnología y de las habilidades humanas, son necesarias representaciones dinámicas y escalables visualmente.

Dependiendo de la técnica, y basándonos en diferentes mejoras en las metáforas subyacentes, se pueden agregar ciertas interacciones que, en principio, incrementarán su escalabilidad visual ([6]). A modo de ejemplo se detallan distintas técnicas de visualización paradigmáticas junto con intentos de aumentar su escalabilidad:

- *Gráficos de barras*: los atributos se pueden representar con la altura y el color de las barras, las barras se pueden apilar para incrementar la cantidad de atributos. Una forma de incrementar la escalabilidad puede ser cambiar dinámicamente a una nueva representación.
- *Landscapes*: los atributos se mapean a la altura, la forma y el color de glifos dispuestos en una grilla. El problema de la oclusión entre los elementos visuales se resuelve con el uso de transparencias u ordenamiento, entre otras posibilidades.
- *Visualizaciones de grafos*: se pueden representar atributos con el tamaño, el color y la forma del nodo, y en el ancho, el color y el patrón de los arcos que unen los nodos. La escalabilidad disminuye drásticamente a medida que la conectividad del grafo aumenta y los diagramados que minimizan la superposición priorizan la accesibilidad de los datos más que la estructura del grafo.
- *Scatterplots*: la escalabilidad disminuye con grandes cantidades de puntos debido a la superposición. Con animaciones e interacciones se ha logrado incrementar la escalabilidad visual ([4]).

En cuanto a las interacciones que incrementan la escalabilidad ([1]) se encuentran:

- *zoom y panning*
- *multi-resolución*: brindando en forma progresiva menos detalle a medida que la cantidad de datos aumenta,
- *selección*: permitiendo seleccionar subconjuntos de datos para reducir la complejidad visual,
- *etiquetado*: aplicando algoritmos de etiquetado que eviten el solapamiento y usen tamaños de fuente variables.

3. Objetivos de la investigación

Muchos conjuntos de datos pueden ser representados como grafos (por ejemplo redes sociales, tráfico en Internet, hipervínculos en Internet), donde los *nodos* representan las entidades objeto de estudio, y los *arcos*, relaciones entre estas entidades.

Uno de los problemas existentes en visualización de grandes grafos es la poca escalabilidad de los sistemas ([3]): a medida que crece la cantidad de nodos y enlaces las aplicaciones tienden a mostrar sólo la topología del grafo. Además, pocos sistemas explotan el uso de interacciones en estas visualizaciones ([5]).

El principal objetivo de este trabajo de investigación será explorar y definir estrategias e interacciones adecuadas que permitan aumentar la escalabilidad visual en la visualización de grafos. Este objetivo, además, nos obligará a definir métricas y metodologías que permitan evaluar la escalabilidad visual e incursionar en técnicas escalabilidad de software que se ajusten a nuestras necesidades.

Referencias

- [1] Stephen G. Eick and Alan F. Karr. Visual scalability. Technical report, National Institute of Statistical Sciences, 2000.
- [2] Charles D. Hansen and Christopher R. Johnson, editors. *The Visualization Handbook*. Elsevier Academic Press, 2005.
- [3] Ivan Herman, Guy Melançon, and M. Scott Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, 2000.
- [4] Catherine Plaisant Jean-Daniel Fekete. Interactive information visualization of a million items. In *INFOVIS '02: Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, pages 117–124. IEEE Computer Society, 2002.
- [5] Martín L. Larrea. Diagramado esférico. Master's thesis, Universidad Nacional del Sur, 2006.
- [6] James J. Thomas and Kristin A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.
- [7] Beth Ann Yost. *The Visual Scalability of Integrated and Multiple Views for High Resolution Displays*. PhD thesis, Virginia Polytechnic Institute and State University, 2005.