

Towards Data-driven, Context-Sensitive Visual Social Network Analysis

Jukka Huhtamäki, Jaakko Salonen and Jarno Marttila, Hypermedia Laboratory at Tampere University of Technology

The use of graphic images to represent social configurations within a given social system is important because "[i]t allows investigators to gain new insights into the patterning of social connections, and it helps investigators to communicate their results to others" (Freeman, 2009). The era of Web 2.0 and social media sets up a context that seems perfect for conducting *social network analysis* (SNA), a research field that has been developed since 1930's to support the study of the structure of networks of social actors. In Web 2.0, a plethora of potentially inter-connectable Web services is used for digitally mediated social interaction. Digital medium means that the interactions are recorded and thus serve, in principle, as an excellent data source for social network analysis. Further, Web feeds and Web APIs are in the core of Web 2.0 paradigm, enabling, at best, the "serendipitous reuse" (Vinoski, 2008) of the resources in different contexts and applications possibly unthinkable to the publisher of the data.

In practice, however, a lot of problems still hinder the development of social network analysis applications sharing light e.g. to the connections between wiki co-creators, actors of social networking platform or some other digital ecosystem. Due to various reasons, not all the systems provide easy access to the data within the system. In addition, several non-interoperable *de facto* formats are used to represent the data both in the source systems and the analysis tool. Despite the recent rapid development of visual tools for social network analysis including Commetrix, NodeXL, SocialAction, Vizster and others, there is no single tool that can be used to collect, manage and visualise the SNA data - an object envisioned by Freeman (2000).

This all implies that the manual work is needed to conduct the analysis. Instead of being able to concentrate to the analysis process per se, the analyst is often forced to solve problems e.g. related to data formatting or even character encoding. Moving the data between the different tools is often done with various copy-and-paste operations that, when the skills of the analyst allow, are partly automated with sketchy scripts and batch processes. Many useful visualisation tools, particularly the ones falling to the category of Web mashups, require the exposure of the SNA data that is often sensitive by nature. Further, the analysis process is usually separate from the normal workflow. In this article, we are going to describe an approach that can be used to streamline the social analysis process.

We approach social network analysis from data-driven *information visualisation* point of view and aim to close the gap between a) the systems where the SNA data originates and b) the tools that are used to conduct the analysis, thus we refer to *data-driven, context-sensitive visual social network analysis*. Historian Alfred Crosby (1997) has highlighted the power of visualisation by stating that visualisation and measurement are the two main factors enabling the explosive development of modern science (cf. Freeman, 2000). The underlying objective of information visualisation is to serve as an amplifier of the cognition of a user through expressive views giving insight on a certain phenomena represented by the data (Ware, 2004). *Interactive* information visualisation "allows the user to implicitly form mental models of the correlations and relationships in the data, through recognition of patterns, marking or focusing in on those patterns, forming mental hypotheses and testing them, and so on" (Kosara, Hauser ja Gresh, 2003).

With context-sensitivity, we refer to the practice of integrating the visualisations or their launching mechanisms to arbitrary Web applications in the spirit of *augmented browsing*, yet making the visualisation sensible to their context. See Salonen and Huhtamäki (2010) for a detailed description of the means to automatically collect context data when launching visualisations and to inject the visualisation tools to the source applications.

In this article, we are going to explore the means to streamline the development of visual social network analysis applications that are able to aggregate SNA data from wikis, social networking engines, Web APIs and other sources to form sets of *linked SNA data* (cf. linkeddata.org), instantly visualise the data with mashup-like solutions that are based on Web technologies or, alternatively, export the data for further analysis to existing high-end tools. Further, means to attach the visualisations and analysis functionalities to the systems where the data originates from are described in the context of our work. Importantly, we take into account the often sensitive nature throughout the proposed solution.

References

- Crosby, A. W. (1997). *The Measure of Reality: Quantification in Western Europe, 1250-1600*. Cambridge University Press.
- Freeman, L. C. (2000). Visualizing Social Networks. *Journal of Social Structure*, 1(1), [np].
- Freeman, L. C. (2009). Methods of Social Network Visualization. In R. A. Meyers (ed.) *Encyclopedia of Complexity and Systems Science*. Berlin: Springer.
- Kosara, R., Hauser, H., & Gresh, D. (2003). An Interaction View on Information Visualization. In *State-of-the-Art Proceedings of EUROGRAPHICS 2003* (pp. 123–137).
- Salonen, J., & Huhtamäki, J. (2010). Launching Context-Aware Visualisations. In *Proceedings of 3rd International OPAALS Conference on Digital Ecosystems, Aracaju, Brazil, 22- 23 March 2010*. Berlin Heidelberg: Springer.
- Ware, C. (2004). *Information Visualization: Perception for Design* (2nd ed.). San Francisco, CA, USA: Elsevier.
- Vinoski, S. (2008). Serendipitous Reuse. *Internet Computing*, (January/February 2008), pp. 84-87.