

Challenger, a New Way to Visualize Data

Remus Zelina¹, Sebastian Bota¹, Siebren Houtman²,
Jaap Jan van Assen², and Bas Hattink²

¹ Meurs HRM, Baia Mare, Romania
{*rzelina, seby*}@meurs.ro

² Meurs HRM, Woerden, The Netherlands
{*s.houtman, j.assen, b.hattink*}@meurshrm.nl

Abstract. Challenger is a software product that provides fast and online data visualization. This is done by visualizing data (graphs) as a network. Both force based and modularization algorithms are used and experimented with. Challenger facilitates fast and easy understanding of complex data. This is not only a matter of showing one ‘perfect’ visualization, but rather of letting users browse, analyze and ‘play’ visually with (subsets of) data interactively.

Keywords: Challenger, Force-Directed, Clustered Graph.

1 Introduction

This abstract¹ presents Challenger, a software product, originally designed to visualize organizations. Organizations were considered as a network of employees, customers and projects (nodes) and links (i.e. ‘working on project x’) between them (edges). Visualizing makes it easier to understand and interpret the (relations in) organizations. However, as networks grow bigger, it turned out to be complex to draw a meaningful visualization within a reasonable amount of time. Challenger as it is now, is the result of our research on data visualization and our attempt to put this knowledge in a useful and fast application. The solution is useful in general to visualize complex data.

2 Implementation

We use different versions of force based algorithms [1], [2], [3] and modularization algorithms [4], [5], [6]. We split the data (graph nodes) into “almost cliques” by maximizing a global criterion. An almost clique is a subset of nodes of the graph that has many internal links and less external links. The global criterion is a value referring to a certain collection of almost cliques that covers the entire graph and it's

¹ This abstract accompanies a poster that can be obtained by contacting us at <http://www.meurs.ro/challenger>. Here one can also find additional information and demonstrations of Challenger.

formula puts in evidence the above phenomenon. The best known formula for this criterion is the Modularity (Q) [4]. It is defined to be the fraction of edges that fall within the given cliques minus the fraction of edges that are expected to fall within the clique if the edges were distributed at random. Splitting a graph into almost cliques leads to information loss (due to external links). To reduce this we consider influences of all almost cliques over a given node, not only those of the almost clique that contains the given node. To optimize speed, we experimented with and optimized different algorithms in different combinations and architectures. The application is web based, and at the moment we achieve performances about 250.000 nodes and 1.000.000 edges in less than 60 seconds including data transfer.

More important than showing a single visualization, Challenger offers possibilities to visually browse, analyze and redraw the graph. This improves the usability. Challenger offers possibilities to:

- mix a number of entity types (nodes) and link types (edges) and combine them in a single graph;
- involve characteristics of elements in the graph (i.e. gender when persons are the elements) by entering the values (i.e. male; female) as ‘virtual’ nodes and put edges - when relevant - between them;
- filter, search and redraw based on characteristics of nodes as well as weights of edges;
- further analyze specific modules;
- create sub graphs and visualizations based on modules, selections of nodes or 1st and 2nd degree relations of selected nodes;
- zoom in and zoom out and keep track in a mini map of the total graph;
- toggle visualization modes.

References

1. Fruchterman, T., Reingold, E.: Graph Drawing by Force-directed Placement. *Software-Practice and Experience* 21(11), 1129–1164 (1991)
2. Kamada, T., Kawai, S.: An algorithm for drawing general undirected graphs. *Information Processing Letters* 31(1), 7–15 (1989)
3. Noack, A.: Energy Models for Graph Clustering. *Journal of Graph Algorithms and Applications* 11(2), 453–480 (2007)
4. Clauset, A., Newman, M.E.J., Moore, C.: Finding community structure in very large networks. *Physical Review E* 70, 066111 (2004)
5. Wakita, K., Tsurumi, T.: Finding community structure in mega-scale social networks. In: *Proceedings of the 16th International Conference on World Wide Web, Banff, Alberta, Canada*, pp. 1275–1276. ACM, New York (2007)
6. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of community hierarchies in large networks. *J. Stat. Mech.* (10), P10008 (2008)