ExplodeLayout: Comprehending Patient Subgroups in Large Networks

Bryant Dang BS[1], Joseph Mathew MS[1], Tianlong Chen BS[2], Suresh K. Bhavnani PhD[1]
[1]Institute for Translational Sciences, University of Texas Medical Branch, Galveston, TX;
[2]Department of Physics, University of Houston, Houston, TX

Abstract

Networks have been used to successfully identify and comprehend patient subgroups based on their characteristics (e.g., comorbidities or genes) with the goal of designing targeted interventions, a cornerstone of precision medicine. However, current network layout algorithms often fail to reveal patterns in large and dense networks despite having significant clustering. We therefore developed an algorithm called *ExplodeLayout*, which exploits the existence of clusters to automatically "explode" a traditional network layout with the goal of separating overlapping clusters to enhance their comprehensibility. We demonstrate the use of this algorithm to visualize a large dataset extracted from Medicare consisting of readmitted hip-fracture patients and their comorbidities, which enabled clinicians to comprehend patient subgroups and their comorbidities, and to infer mechanisms related to hospital readmission.

Introduction

Although network visualizations have enabled researchers to identify and comprehend patient subgroups based on their characteristics (e.g., genes or comorbidities),[1] current layout algorithms such as *Fruchterman–Reingold* (FR) often fail to reveal such subgroups despite the networks having significant clustering. For example, Fig. 1A shows a bipartite network consisting of all 30-day readmitted hip fracture patients (n=6150) extracted from the 2010 Medicare database that had at least one of the 8 significant comorbidities shown. Unfortunately, the network layout generated by FR is difficult to comprehend due the overlapping clusters despite (1) the network having strong clusteredness (co-clustering modularity[2]=0.440) that was significant (p<.001) compared to 1000 random permutations of the network, and (2) the use of distinct colors to denote each co-cluster of patients and comorbidities. Here we demonstrate *ExplodeLayout* an approach that exploits information about such clusters to enhance their comprehensibility.

Method and Results

ExplodeLayout is implemented in R, and requires as input (1) coordinates of network nodes generated by a traditional force-directed algorithm such as FR, (2) the number and members of co-clusters generated by a modularity algorithm[1] (standard in most network applications), and (3) a user-selected circle radius using a scroll bar to determine the distance by which the node clusters are moved. The algorithm uses these inputs to (1) generate an imaginary circle for the given radius with n equidistant points, where n=number of the clusters, (2) move the nodes within each cluster by the same distance and angle such that the centroid of that cluster coincides with one of the points on the circle. As shown in Figure 1B, this transformation effectively "explodes" the cluster locations by preserving the distances among nodes within the clusters, and by *not* preserving the distances among nodes across the clusters. This trade-off between not preserving inter-cluster node distances and having greater separation between the clusters enables a researcher to comprehend the relative size and topology of each cluster. An optimal layout is generated iteratively by the researcher to achieve a desired balance between cluster separation and layout compactness.

Layouts generated from *ExplodeLayout* have enabled important insights[1] into the risk factors for subgroups of readmitted hip-fracture patients, and therefore could be an important complement to force-directed algorithms such as FR. However, for networks with a large number of clusters, the user would need to select a large circle radius to enable cluster separation, resulting in unused space in the center of the layout. We therefore implemented the ability to specify the number of nested circles enabling the central region of the layout to be filled and therefore still maintaining separation with compactness. In future research we plan to automatically suggest to the user the appropriate number of nested circles, which can be modified and therefore enable the rapid generation of compact and comprehensible layouts consisting of hundreds of thousands of nodes, and with many patient subgroups.

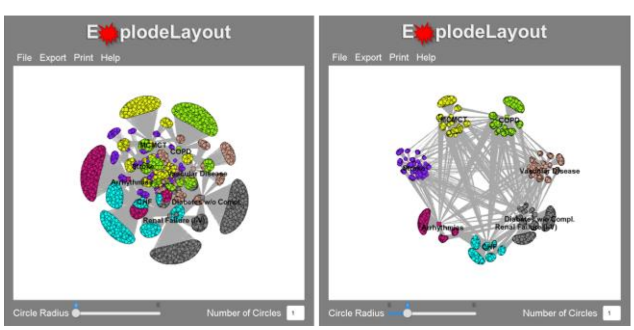

Figure 1. *ExplodeLayout* exploits clustering of patients and their characteristics in dense networks (A) by "exploding" the cluster locations based on user input (B), to reduce overlap and enhance their comprehensibility.

References

1. Bhavnani SK, et al. How Comorbidities Co-occur in Readmitted Hip Fracture Patients: From Bipartite Networks to Insights for Post-Discharge Planning. Proceedings of AMIA Summit on CRI (2015).

2. Newman M. Networks: An Introduction. Oxford University Press; 2010.