Exploring the Distribution of Local Neighborhood Structures in Large Networks

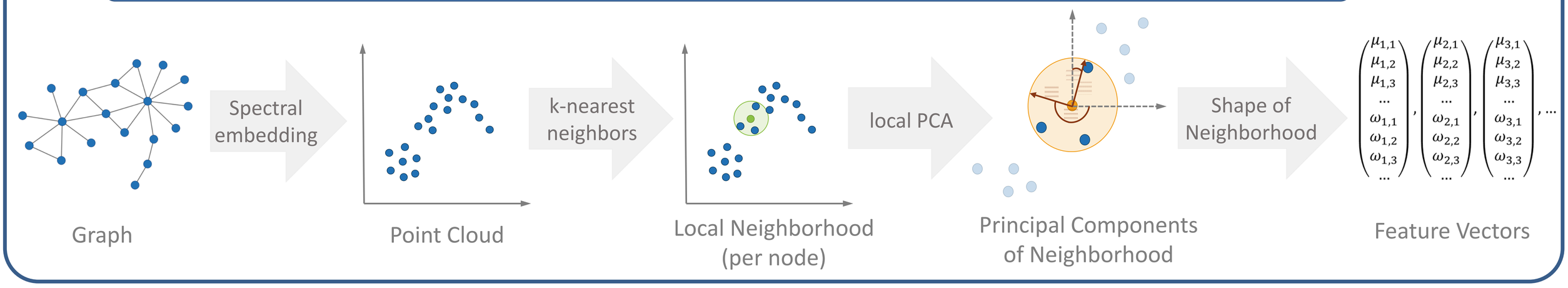
Shenghui Cheng[2], Claudia Dahl[1], Joachim Giesen[1], Philipp Lucas[1], Klaus Mueller[2]

[1] Friedrich-Schiller-University Jena, Germany
[2] Visual Analytics and Imaging Lab, Stony Brook University, USA and SUNY, Korea

Introduction

Network exploration techniques aim at identifying structures like clusters, central nodes or motifs in networks. We introduce a new type of local neighborhood feature that is spectrally defined, i.e., based on the eigenvalues and eigenvectors of some matrix associated with the network, and show that well established visual analytics techniques for exploring high-dimensional point clouds provide an effective means for the exploration of the distribution of these neighborhood structures. Experiments on real world social networks demonstrate that our approach is indeed capable of revealing interesting neighborhood structures that are not easily accessible otherwise.

Extraction Pipeline of Local Neighborhood Features Vectors



Net² Tool

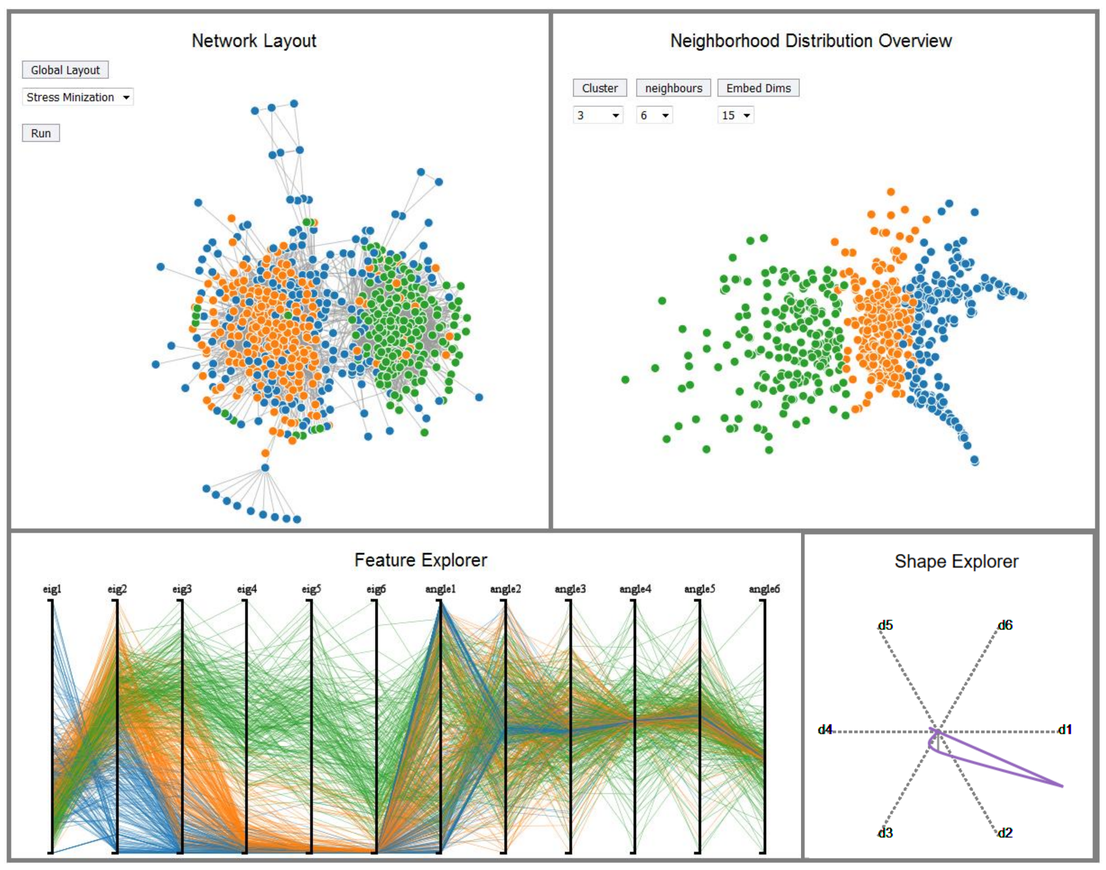


Figure 1: Screenshot of the Net² tool, with the TWITTER data set loaded.

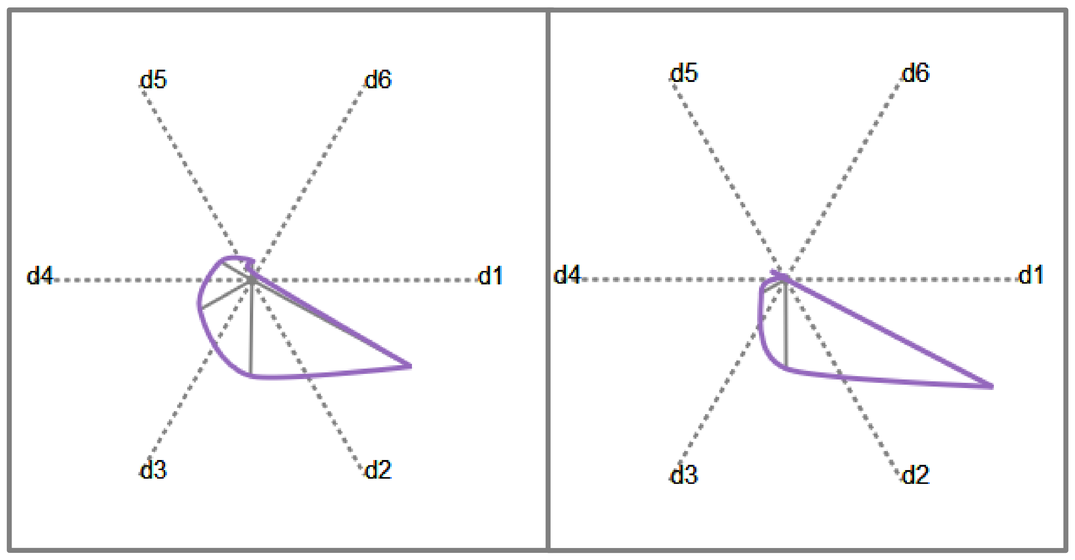


Figure 2: ShapeExplorer views that summarize the geometry of the local neighborhood structure for both the dominant groups in the TWITTER n*etwork.*

Net² is the interactive local neighborhood explorer tool that we implemented to validate our approach.
It features four fully linked views: (i) the Network Layout view (showing a node-link diagram of the input network), (ii) the Neighborhood Distribution Overview (showing a multidimensional scaling plot of the feature vectors), (iii) the Feature Explorer (showing a parallel coordinate plot of the feature vectors), and (iv) the Shape Explorer (showing the average shape of the neighborhood of the currently selected nodes).
It also allows a variety of interactions to assist network exploration: (i) selecting nodes (linked across all views), (ii) filtering nodes according to feature dimension values, (iii) projections (removing some feature dimensions) and (iv) automatic k-means clustering. Moreover, parameters of the feature extraction pipeline can be changed on the fly.

Case Studies

We demonstrate our method at the example of a real world TWITTER network that captures a political debate. See Figure 1 and Figure 2. Two of our findings are:

1. The two major clusters in the original network correspond to clusters in the feature space. That means the two groups exhibit different local neighborhood structure (see Figure 2) and hence a different communication pattern within the two groups. It is important to note that this difference cannot be seen just from the graph layout.
2. A three-means clustering of the feature vectors finds three meaningful groups of nodes: the two main clusters (green and orange, see above), and a group of boundary/interface nodes (blue). Their interface character is confirmed by a look at their spiky neighborhood shape as shown in the ShapeExplorer of Figure 1.

Email: [2] {shecheng, mueller}@cs.sunysb.edu,
[1] {claudia.dahl, joachim.giesen, philipp.lucas}@uni-jena.de

seit 1558

SUNY Korea
Stony Brook University Graduate Degree Programs
The State University of New York

Stony Brook University