Social Status and Role Analysis of Palin's Email Network

Xia Hu and Huan Liu

Computer Science and Engineering, Arizona State University, USA

{xiahu, huan.liu}@asu.edu

Motivation and Significance

- Studying social status and the corresponding social roles in a personal email network can help understand complex interaction patterns in a social network.

- The availability of Sarah Palin's email corpus presents such an opportunity in an email network.

Data Preparation

Step1: Conversion from scanned email exchanges (images) to a text corpus.
Step 2: Name Entity Resolution
Step 3: Email Network Reconstruction

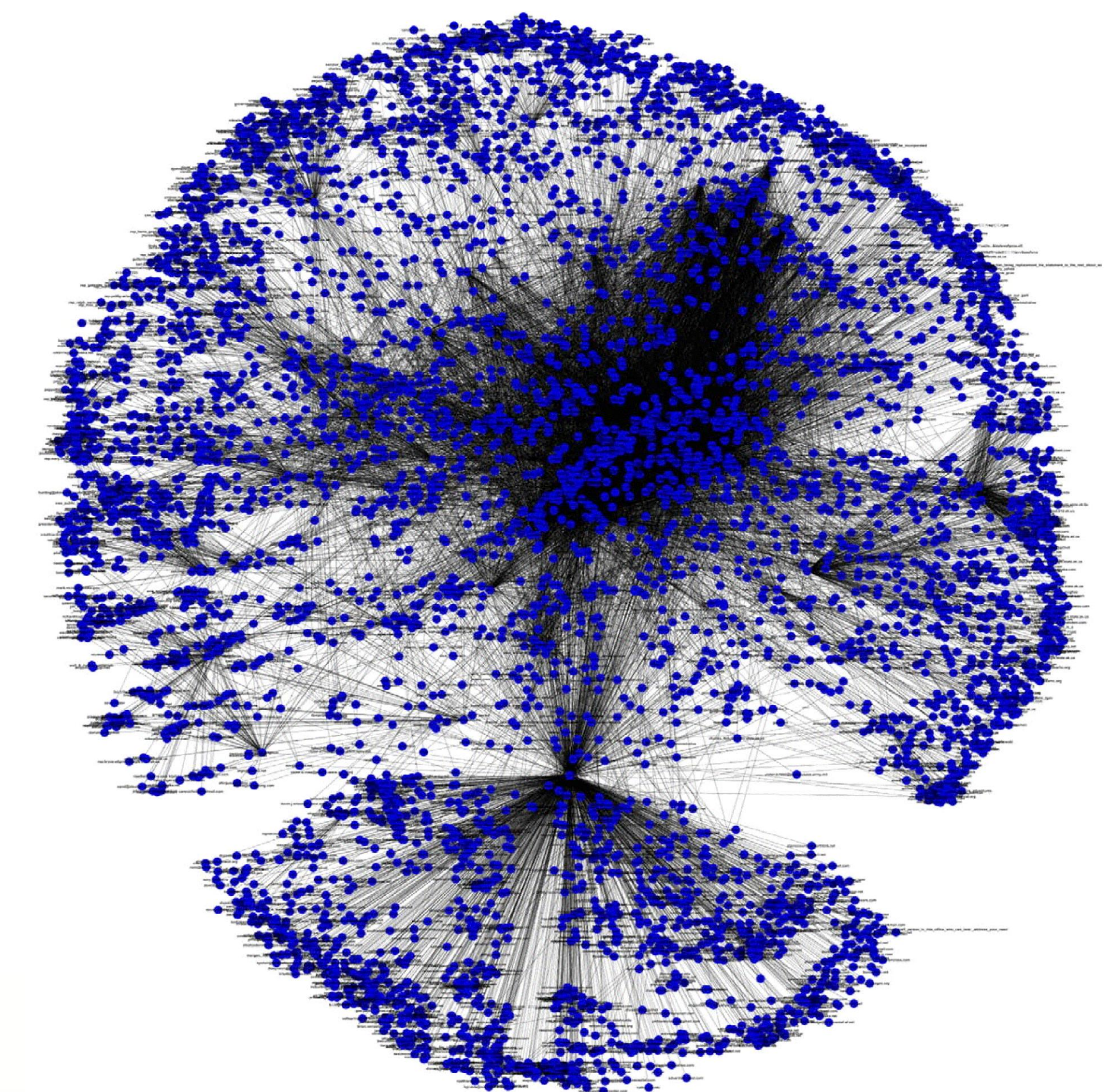※ This dataset is available upon request. Please contact the first author.

Social Status and Role Analysis

- To explore complex social relations of Sarah Palin in her social network, we employ a "knock-out" technique in the experiment.

- Most of the nodes (84.9%) and unique links (72.2%) still exist in the biggest component.

- The clustering coefficient even increases without Palin, which demonstrates that she had a robust structure of supporting staff.

- The removal of Sarah Palin significantly improve the understanding of her email network.

Table 1: Networks With & Without Sarah Palin

	With Palin	Without Palin
# of Nodes	4446	4445
Biggest Component	4446	3773
# of Edges	59589	22177
Clustering Coefficient	0.146	0.171
Network Centralization	0.220	0.072

Visualization

Conclusion

- We employ text analytical tools to reconstruct Palin's email network.

- We provide a preliminary study of Palin's social status and social role in the network.

Acknowledgments

Data Mining and Machine Learning Lab

ARIZONA STATE UNIVERSITY