

---

# vGraph: A Generative Model for Joint Community Detection and Node Representation Learning – Supplementary Material

---

Anonymous Author(s)

Affiliation

Address

email

## 1 A Datasets

2 Citeseer, Cora, Cornell, Texas, Washington, and Wisconsin are available online<sup>1</sup>. For Youtube,  
3 Amazon, and Dblp, we consider subgraphs with the 5 largest ground-truth communities due to the  
4 runtime of the baseline methods.

5 **Facebook**<sup>2</sup> is a set of Facebook ego-networks. It contains 10 different ego-networks with identified  
6 circles. Social circles formed by friends are regarded as ground-truth communities.

7 **Youtube**<sup>3</sup> is a network of social relationships of Youtube users. The vertices represent users; the  
8 edges indicate friendships among the users; the user-defined groups are considered as ground-truth  
9 communities.

10 **Amazon**<sup>4</sup> is collected by crawling amazon website. The vertices represent products and the edges  
11 indicate products frequently purchased together. The ground-truth communities are defined by the  
12 product categories on Amazon.

13 **Dblp**<sup>5</sup> is a co-authorship network from Dblp. The vertices represent researchers and the edges  
14 indicate co-author relationships. Authors who have published in a same journal or conference form a  
15 community.

16 **Coauthor-CS**<sup>6</sup> is a computer science co-authorship network. We chose 21 conferences and group  
17 them into five categories: *Machine Learning*, *Computer Linguistics*, *Programming language*, *Data*  
18 *mining*, and *Database*.

---

<sup>1</sup><https://linqs.soe.ucsc.edu>

<sup>2</sup><https://snap.stanford.edu/data/ego-Facebook.html>

<sup>3</sup><http://snap.stanford.edu/data/com-Youtube.html>

<sup>4</sup><http://snap.stanford.edu/data/com-Amazon.html>

<sup>5</sup><http://snap.stanford.edu/data/com-DBLP.html>

<sup>6</sup><https://aminer.org/aminernetwork>

## 19 B Visualization

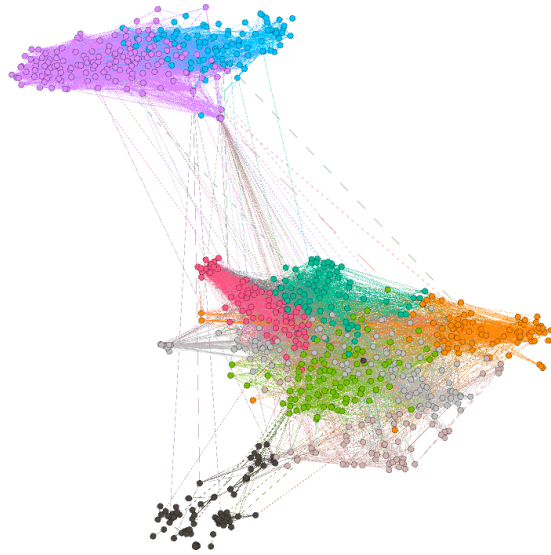


Figure 1: Visualization of the result of vGraph on the facebook1684 dataset. The coordinates of the nodes are determined by t-SNE of the node embeddings.

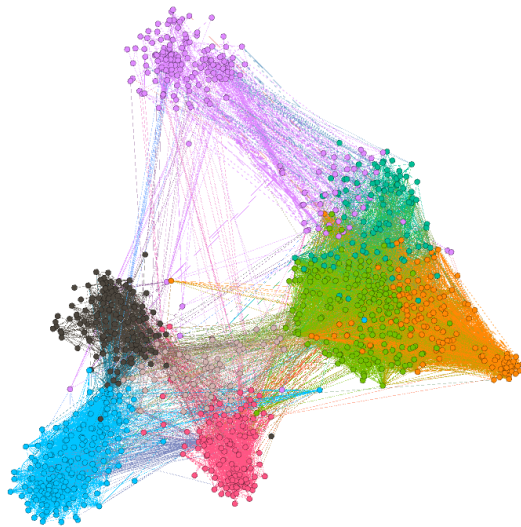


Figure 2: Visualization of the result of vGraph on the facebook107 dataset. The coordinates of the nodes are determined by t-SNE of the node embeddings.

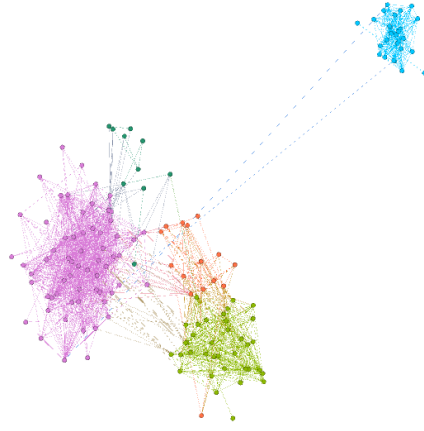


Figure 3: Visualization of the result of vGraph on the facebook414 dataset. The coordinates of the nodes are determined by t-SNE of the node embeddings.

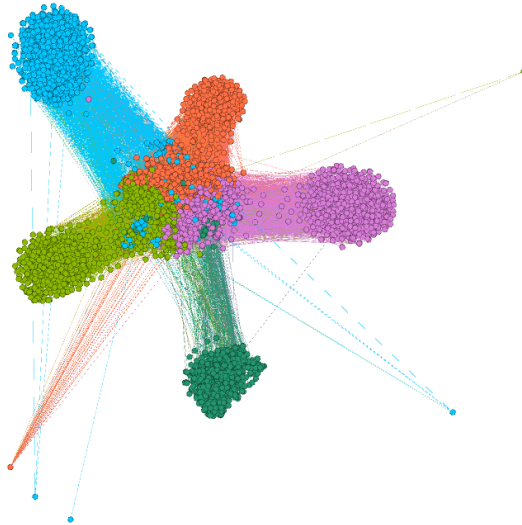


Figure 4: Visualization of the result of vGraph on the Youtube dataset. The coordinates of the nodes are determined by t-SNE of the node embeddings.