By *Jure Leskovec*

STANFORD UNIVERSITY

SNAP for C++
SNAP for Python
SNAP Datasets
BIOSNAP Datasets
What's new
People
Papers
Projects
Citing SNAP
Links
About
Contact us

Open positions

Open research positions in **SNAP** group are available at **undergraduate, graduate** and **postdoctoral** levels.

Graph Embedding with Self Clustering: Deezer, February 13 2018

Dataset information

The data was collected from the music streaming service Deezer (November 2017). These datasets represent friendship networks of users from 3 European countries. Nodes represent the users and edges are the mutual friendships. We reindexed the nodes in order to achieve a certain level of anonimity. The csv files contain the edges -- nodes are indexed from 0. The json files contain the genre preferences of users -- each key is a user id, the genres loved are given as lists. Genre notations are consistent across users. In each dataset users could like 84 distinct genres. Liked genre lists were compiled based on the liked song lists. The countries included are Romania, Croatia and Hungary. For each dataset we listed the number of nodes an edges.

The data was collected in November 2017.

GEMSEC paper: arxiv.org
GEMSEC Project: Github

Country	#Nodes	#Edges
RO	41,773	125,826
HR	54,573	498,202
HU	47,538	222,887

Source (citation)

B. Rozemberczki, R. Davies, R. Sarkar and C. Sutton. GEMSEC: Graph Embedding with Self Clustering. 2018.

```
@inproceedings{rozemberczki2019gemsec,
    title={GEMSEC: Graph Embedding with Self Clustering},
    author={Rozemberczki, Benedek and Davies, Ryan and Sarkar, Rik and Sutton, Charles},
    booktitle={Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Netwo
    pages={65-72},
    year={2019},
    organization={ACM}
}
```

Files

File	Description
gemsec_deezer_dataset.tar.gz	Deezer data from February 13 2018