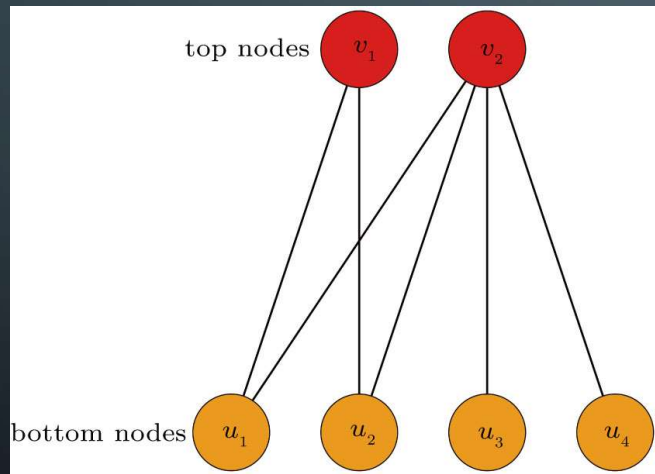


FINAL PROJECT: STUDY OF COMMUNITIES IN GITHUB WITH A MEMBERSHIP NETWORK

AUTHOR: MANUELLE NDAMTANG

NETWORK FEATURES

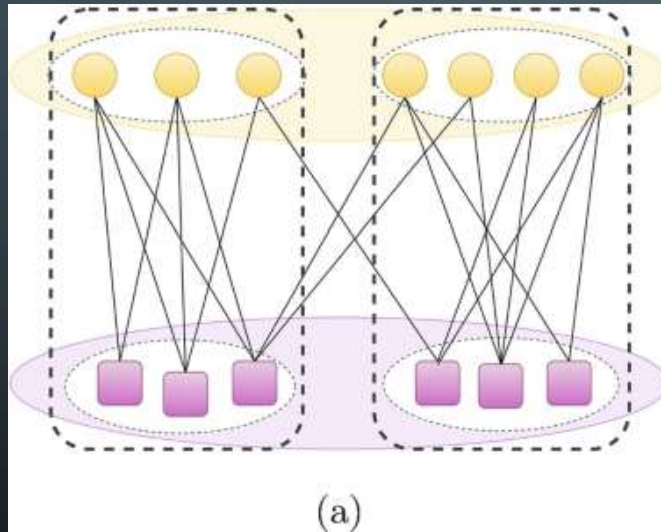
- Bipartite graph



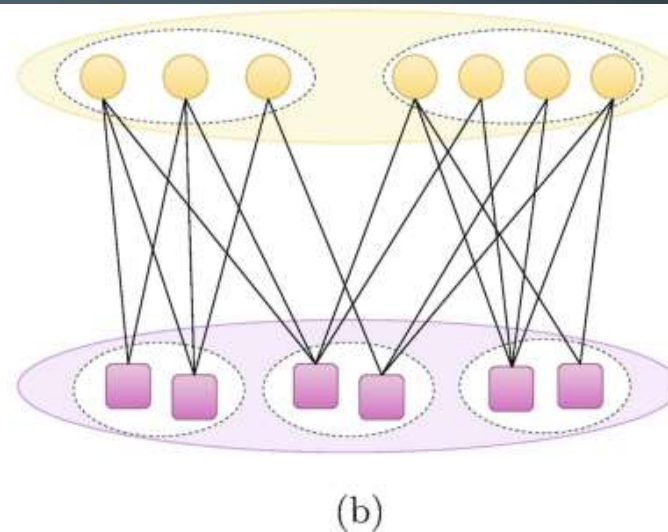
- Number of nodes: 177 386
- Number of edges: 440 237
- The number of connected components: 15 067
- Number of users: 56 519
- Number of projects: 120 867

TWO WAYS TO DEFINE COMMUNITIES

- One-to-one correspondence communities



- Many-to-many correspondence communities



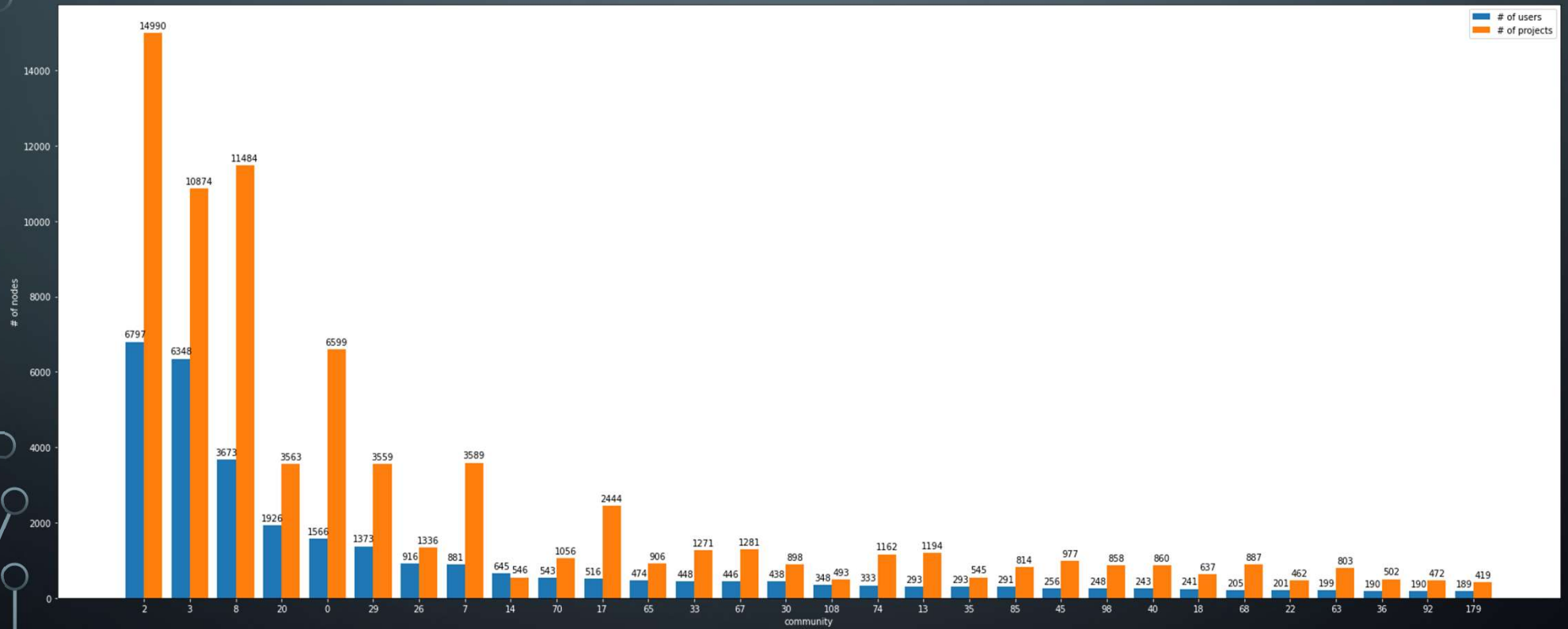
COMMUNITY DETECTION IN THE GRAPH WITH THE BIPARTITE LOUVAIN ALGORITHM

Library from the paper: “Improving Performances of Top-N Recommendations with Co-clustering Method”

- ✓ Easy to use
- ✗ Use projection
- ✗ Instability

The number of communities detected: ~17 690

COMMUNITY DETECTION IN THE GRAPH WITH THE BIPARTITE LOUVAIN ALGORITHM



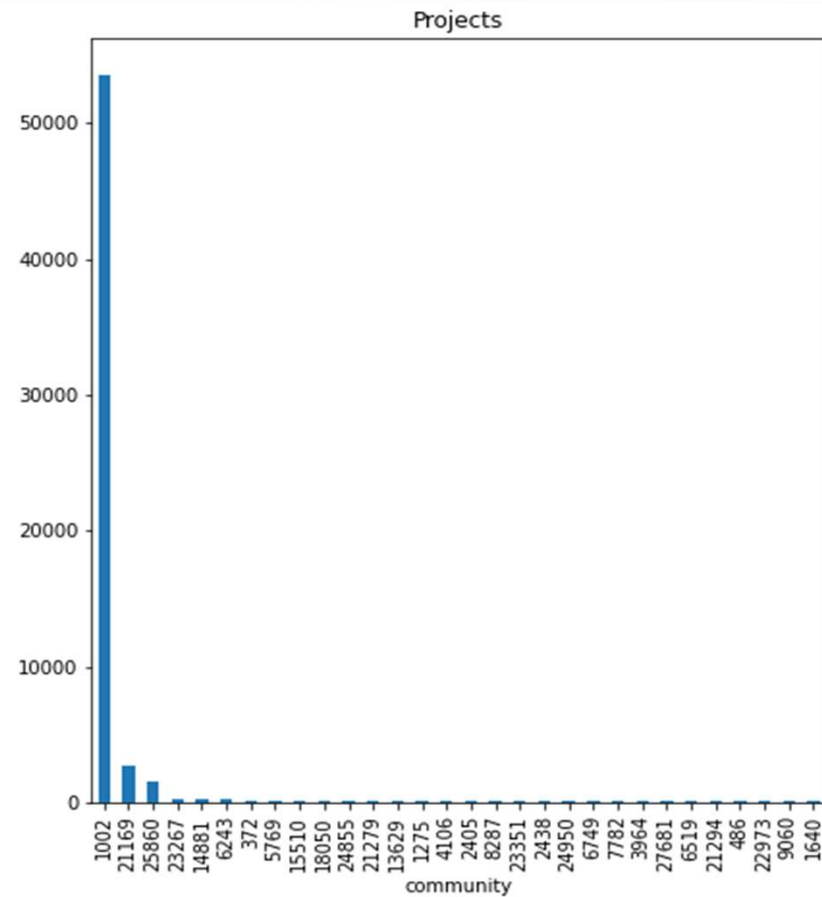
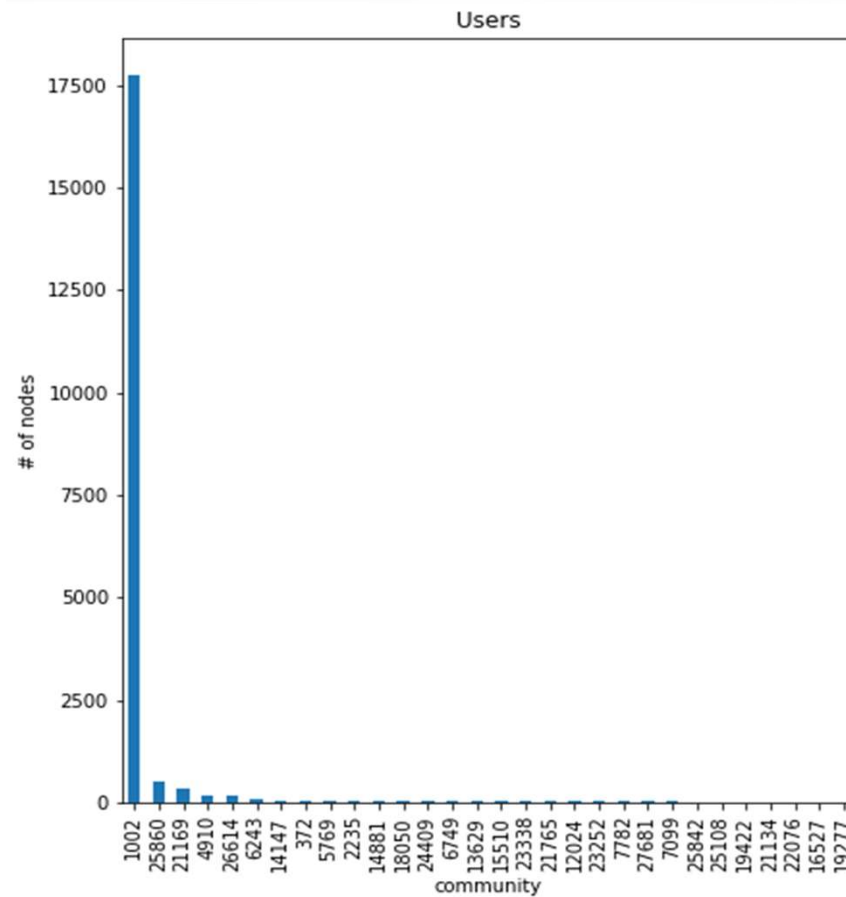
COMMUNITY DETECTION USING BIMLPA

From paper: “BIMLPA: Community Detection in Bipartite Networks by Multi-Label Propagation”

- ✓ More precise
- ✓ Stable algorithm
- ✗ Need of many iterations
- ✗ Hard to apply on large scale graph

After 100 iterations (max), number of communities: (28 558, 27 264)

COMMUNITY DETECTION USING BIMLPA



COMMUNITY PREDICTION: APPLY MACHINE LEARNING TO A SUBGRAPH

- Build model to predict community for a node
- Community based on bipartite Louvain algorithm
- Use graph embedding
- Create a predictive model