



Exploring the
FACEBOOK
community





OVERVIEW

01

Introduction

02

Research questions

03

Dataset

04

Key network statistics

05

Degree distribution

06

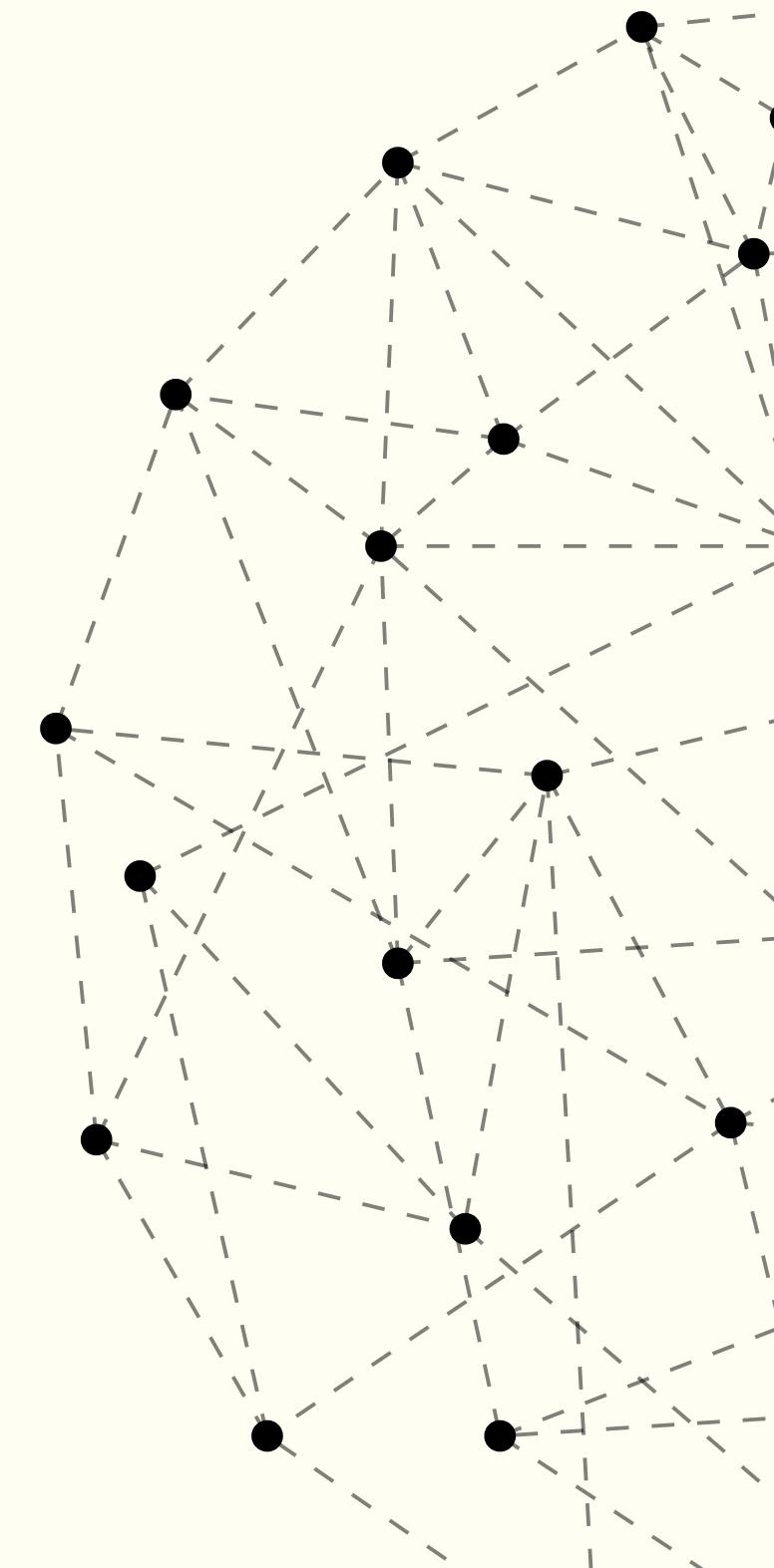
Scale-free network

07

Community detection

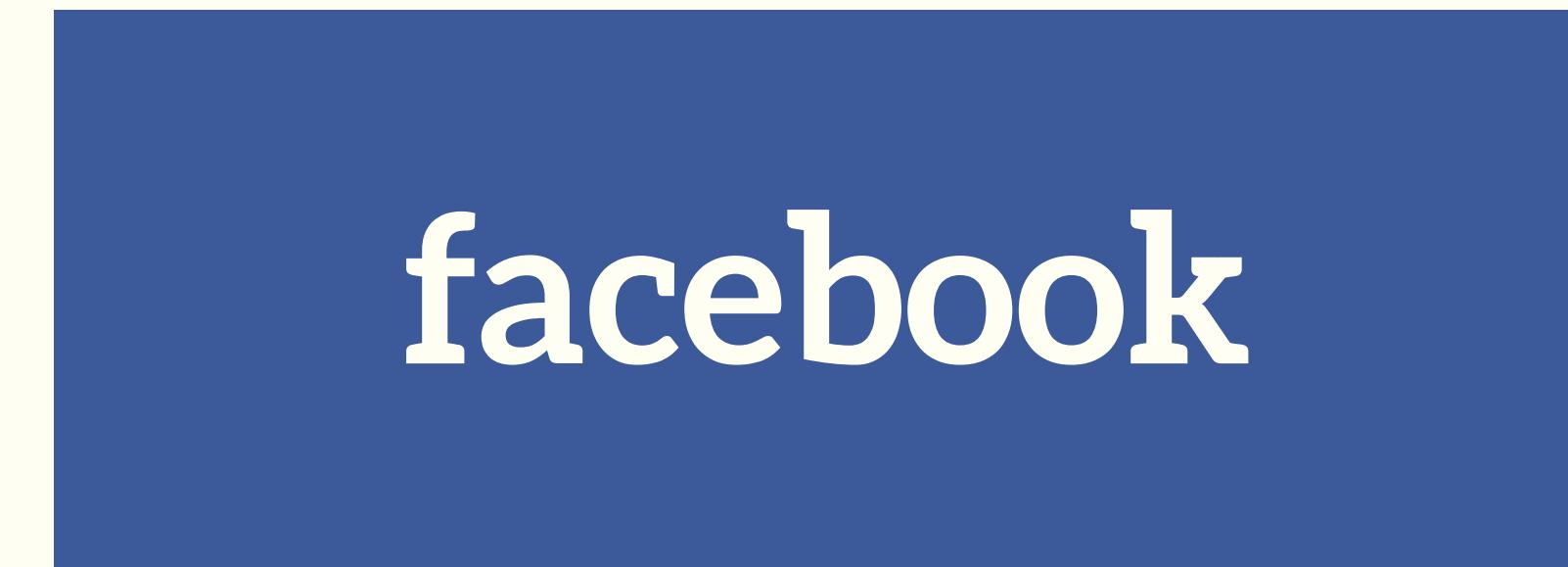
08

Conclusions



Share text, photos with individuals users accepted as "friends" or, using various privacy settings, with the public.

Created in 2004



2017: 1.8 billion active users

Engage in direct communication through Messenger, participate in interest-based groups and stay informed about the activities of Facebook friends and followed pages.

AIM OF THE PROJECT

- **explore the features of THE social network**

I created my Instagram account, which has become my primary social platform, at the end of 2016. Up until then, Facebook was the only social media I used. Specifically, I used Facebook again two years ago when I was searching for a room in Milan. The aspects that work well on Facebook are the events and groups, where interaction is more effective and clear compared to other applications.

- **does it still work as a social network 13 years after its creation?**

Do likes between pages happen randomly? Are the pages still active and engaged? One would expect so, given the substantial user base on this platform, especially considering the significant presence of older generations* using it as a primary source of news.

*The so called baby boomers, individuals born between mid-1946 and mid-1964.



Source: <https://snap.stanford.edu/data/facebook-large-page-page-network.html>

DATASET

The graph was collected through the Facebook Graph API in November 2017.

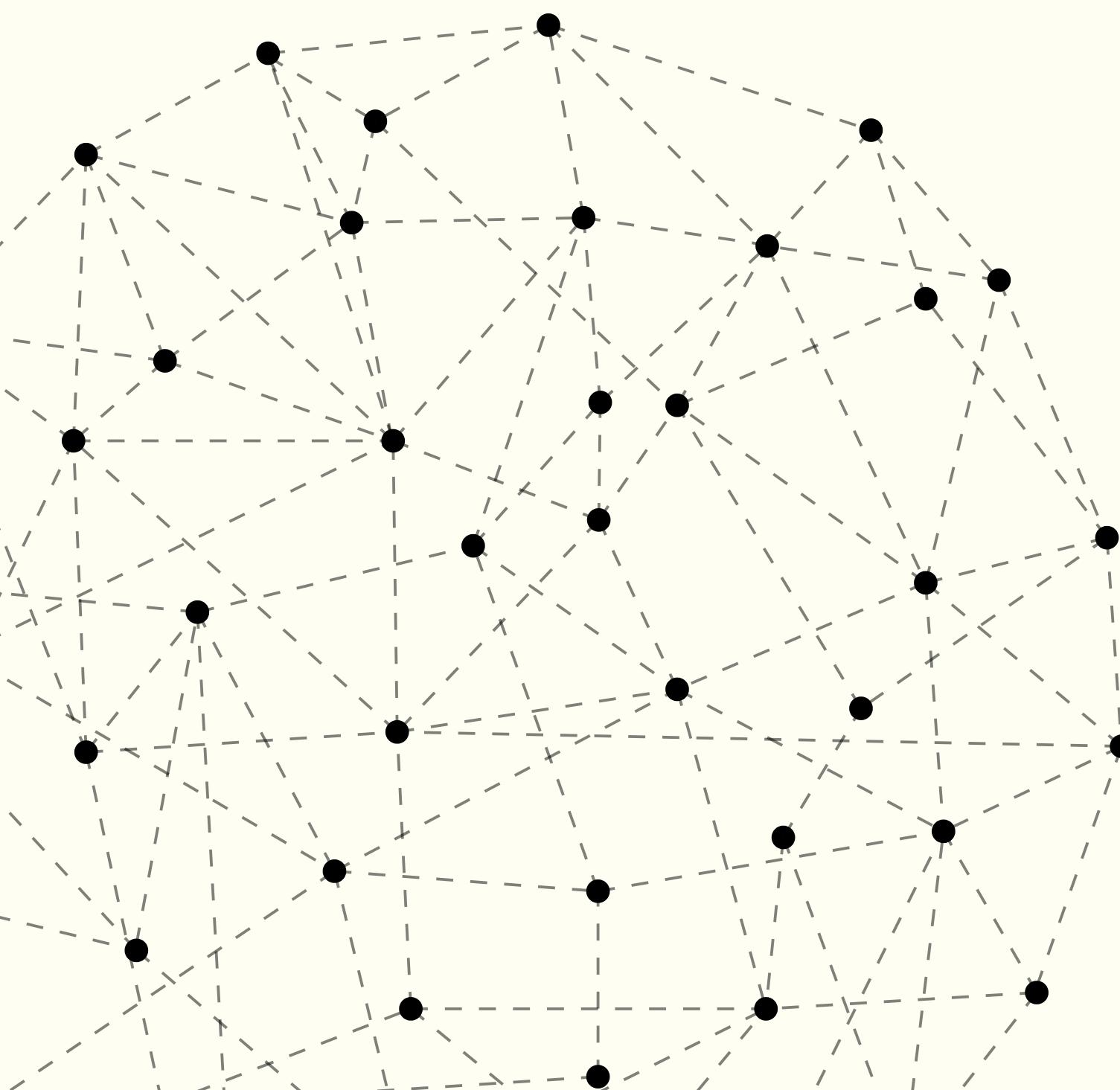


Nodes are official Facebook pages while the links are *mutual likes* between sites.

Attributes: pages are labelled as categories which are defined by Facebook. These categories are: ***politician*** (5768), ***governmental organization*** (6880), ***television show*** (3327) and ***company*** (6495).



Network description



Connected and
undirected graph

Edges:
mutual likes between
Facebook pages
171002 edges

Nodes:

- verified Facebook pages
- 22470 nodes

KEY NETWORK FEATURES

global measure

DENSITY: 0.000677



the graph is sparse: the number of edges is much less than the possible number of edges.

local measure

MEAN DEGREE: 15.22

MEDIAN DEGREE: 7

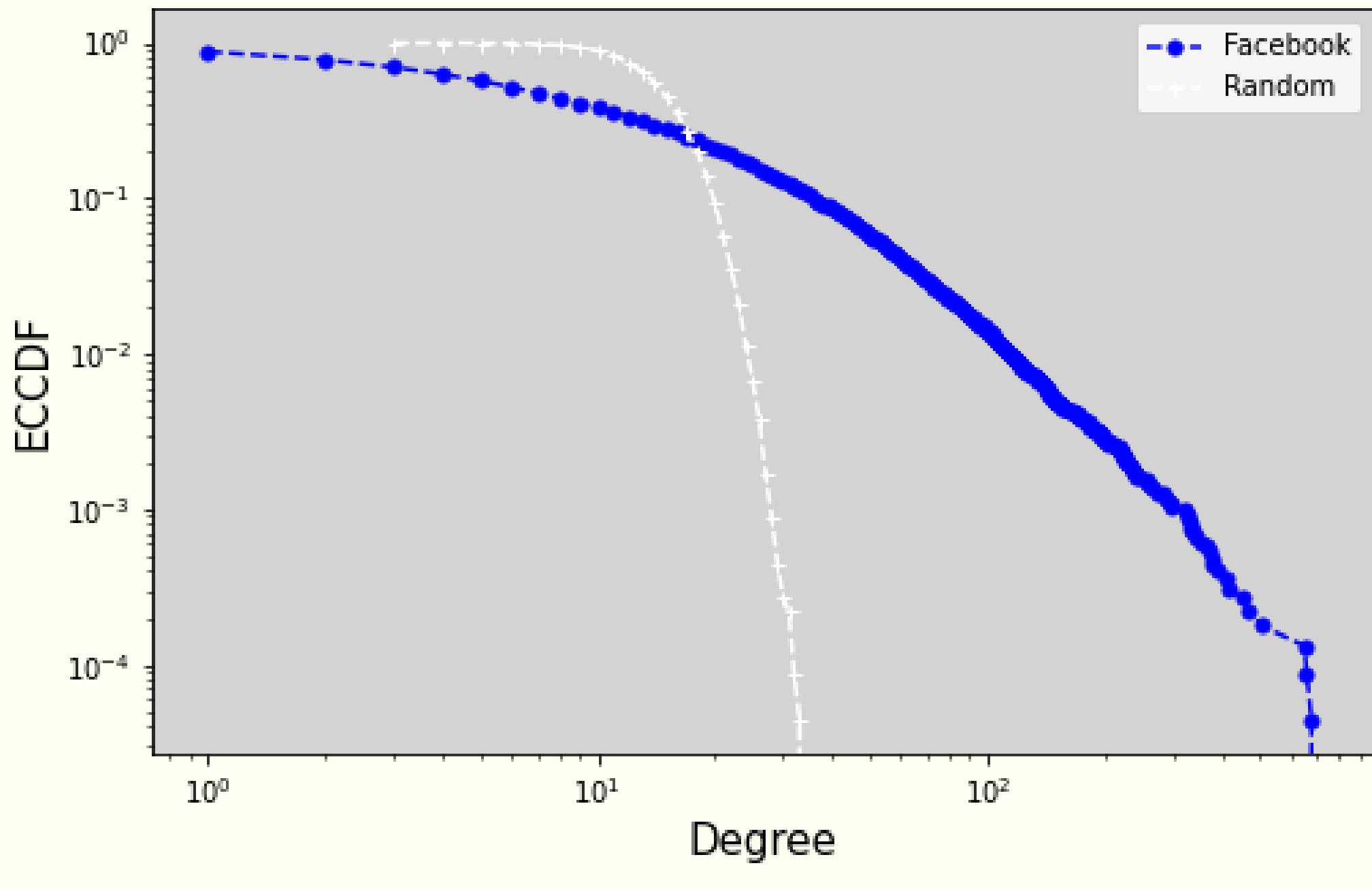
STANDARD DEVIATION: 26.41

RANGE DEGREE: [1, 709]

AVG SHORTEST PATH: 4.97



DEGREE DISTRIBUTION



The ECCDF of the network is quasi linear, meaning that the degree distribution follows a power-law distribution, different with respect to a random network.

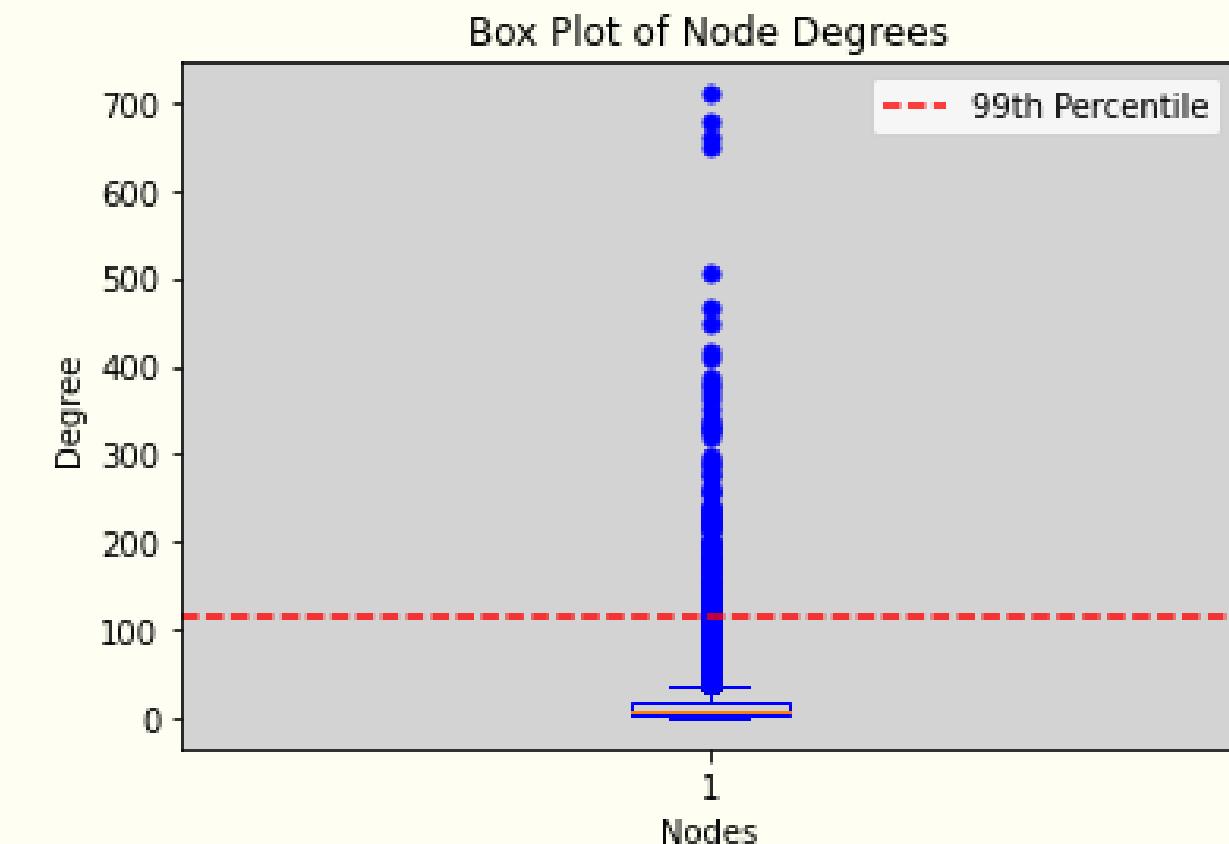
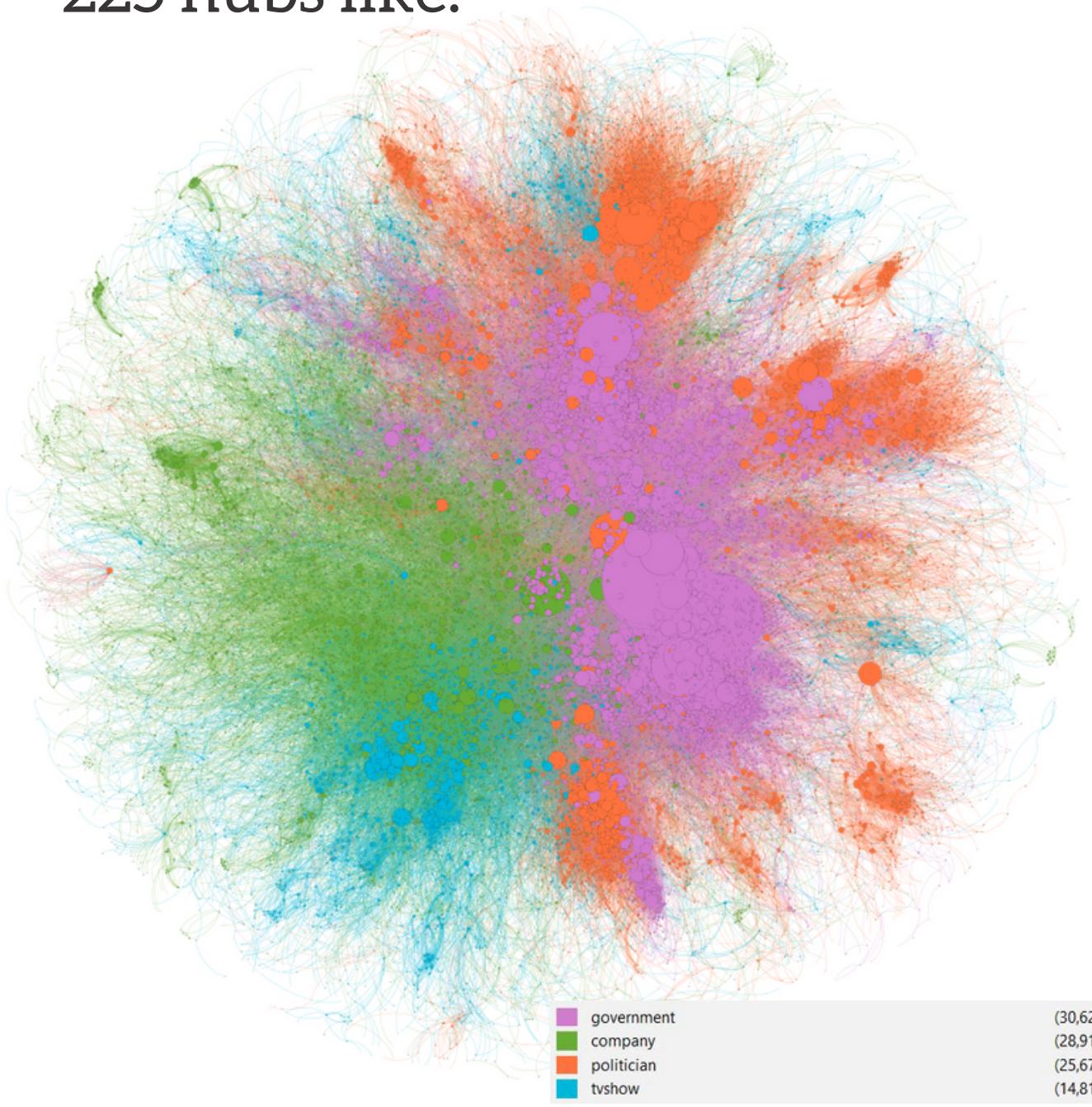
It is a **scale-free network!**



DEGREE DISTRIBUTION

Threshold (considering the 99th percentile) to be a hub: degree > 114.

225 hubs like:



- Bill Shorten MP: Australian politician
- U.S. Forest Service: agency within the United States Department of Agriculture (USDA) that is responsible for managing and protecting the national forests and grasslands of the United States
- Dilma Rousseff: Brazilian economist and politician who served as the 36th president of Brazil
- Marie-Claude Bibeau: Canadian politician who has been involved in federal politics...

ASSORTATIVITY

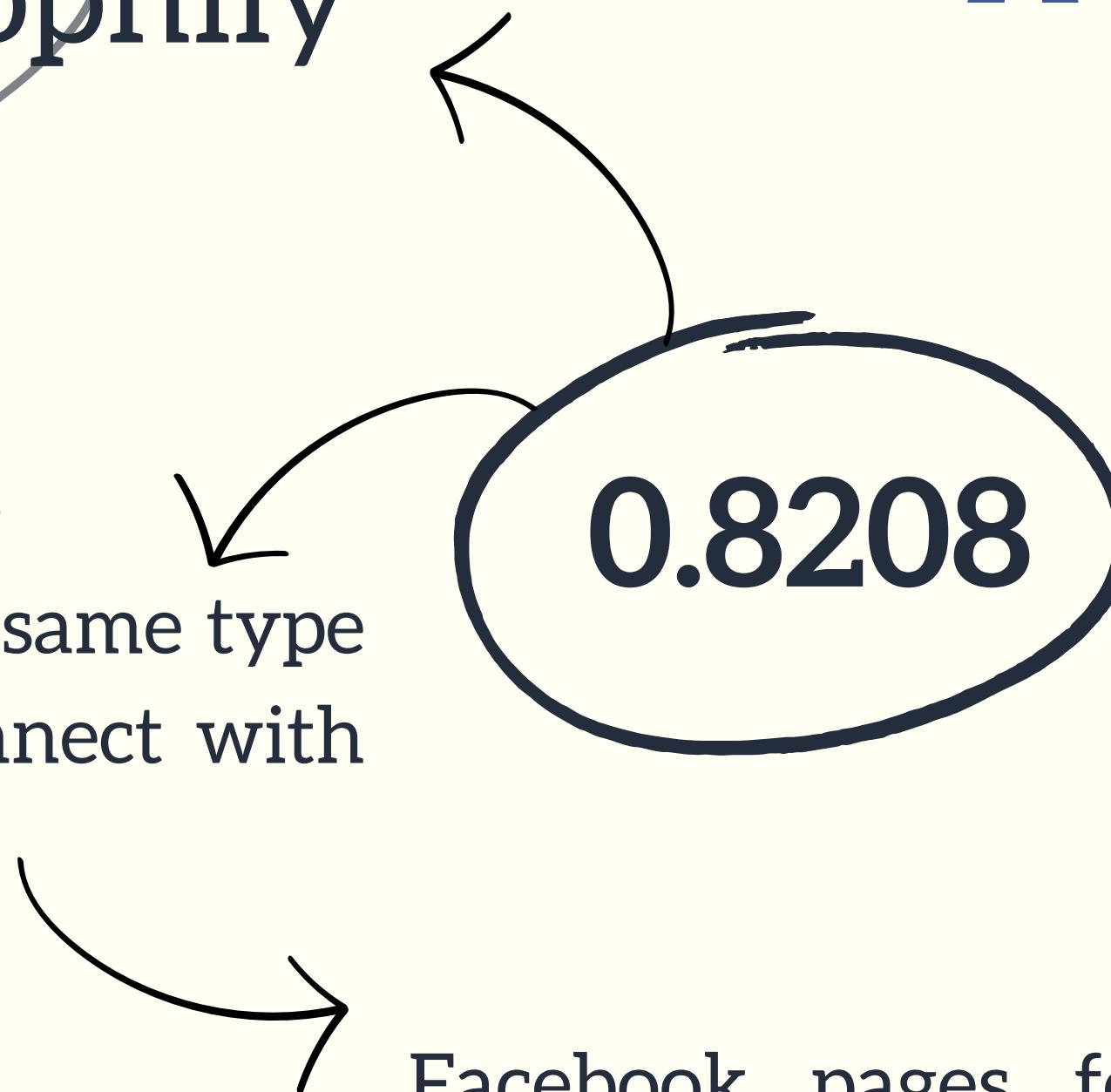
attribute assortativity

Homophily

POSITIVE!

Facebook pages of the same type
are more likely to connect with
each other.

0.8208



Facebook pages fosters a network structure
where pages with similar topic clusters of mutual
support: likes seem to be not randomly!

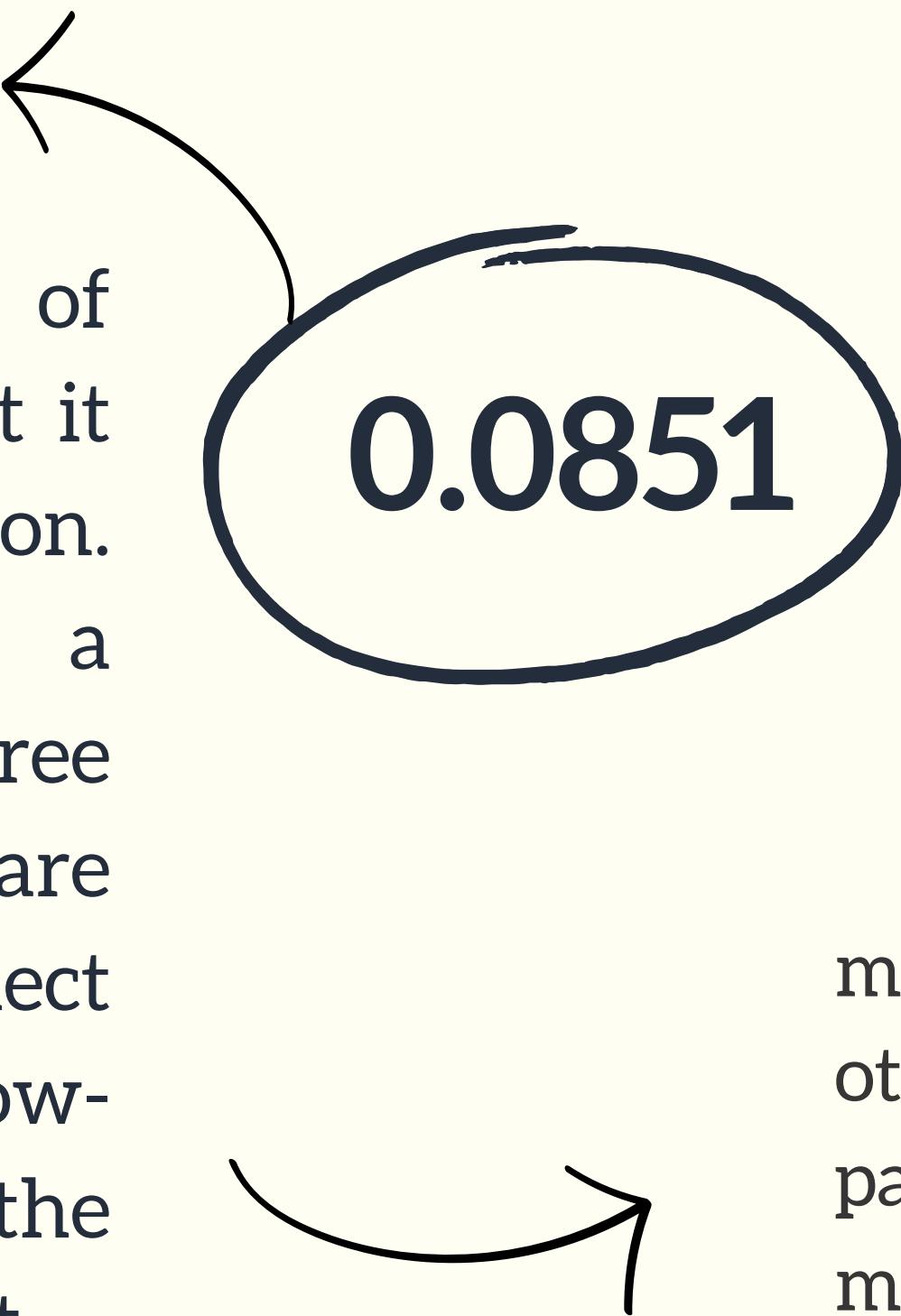


ASSORTATIVITY

degree assortativity

SLIGHTLY
POSITIVE!

Slight tendency for nodes of similar degree to connect, but it is not a very strong correlation. The network might have a structure where high-degree nodes (low-degree nodes) are slightly more likely to connect with other high-degree (low-degree nodes) nodes, but the tendency is not very dominant.



Hypothesis:
more famous pages support each other, but they also support smaller pages that, for example, frequently mention their posts.

COMMUNITY DETECTION

Louvain method

- To find (non overlapping) communities in large networks.
- It is based on modularity maximization.
- Modularity is a measure of the quality of each partition.

Each node is assigned to its own community. For each node, the algorithm evaluates the potential gain in modularity by moving the node to the community of its neighbors. The node is then placed in the community that maximizes this gain (only if the gain is positive). This process is repeated for all nodes sequentially until no further improvement in modularity can be achieved.

1

2

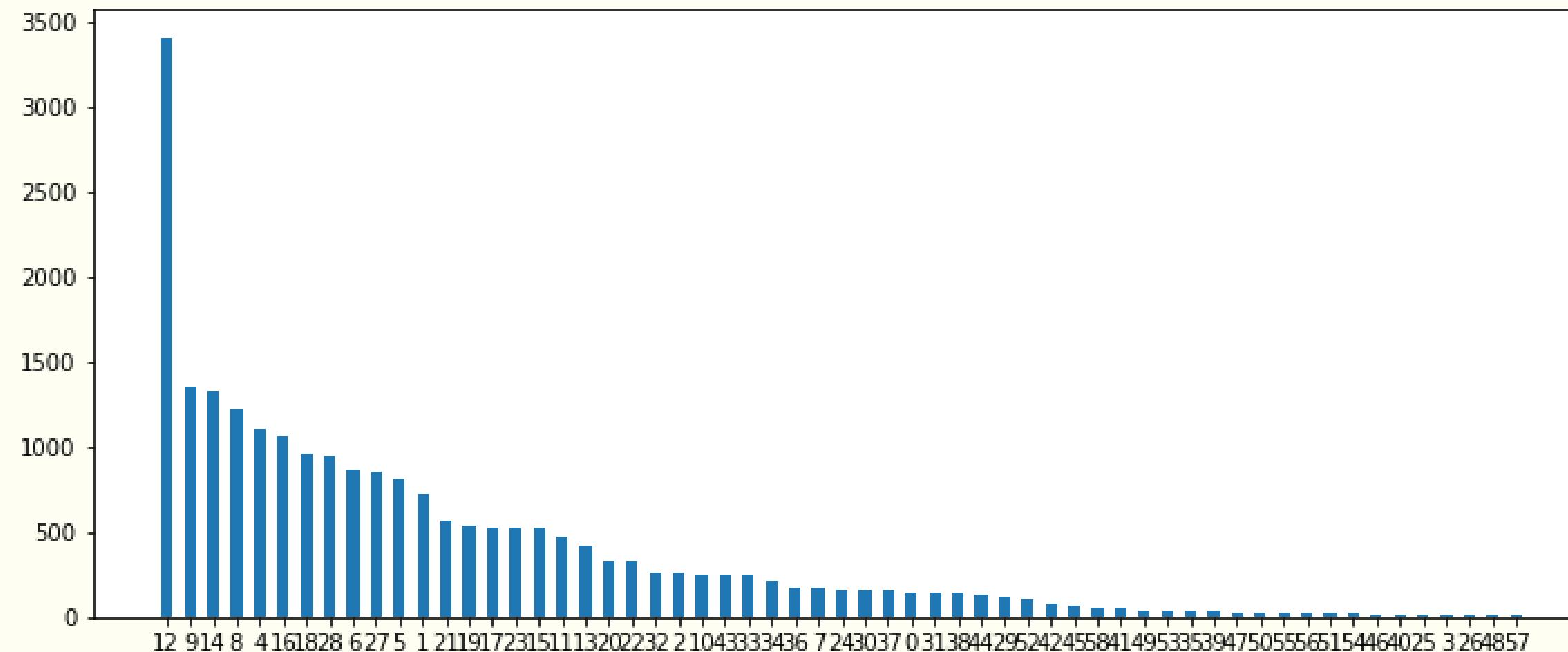
A new network is constructed, where the nodes represent the communities identified in phase 1. The weight of a link between two nodes in this new network is determined by the sum of the weights of the links between nodes in the corresponding communities.

Pass: repetition of phase 1 and 2.

COMMUNITY DETECTION

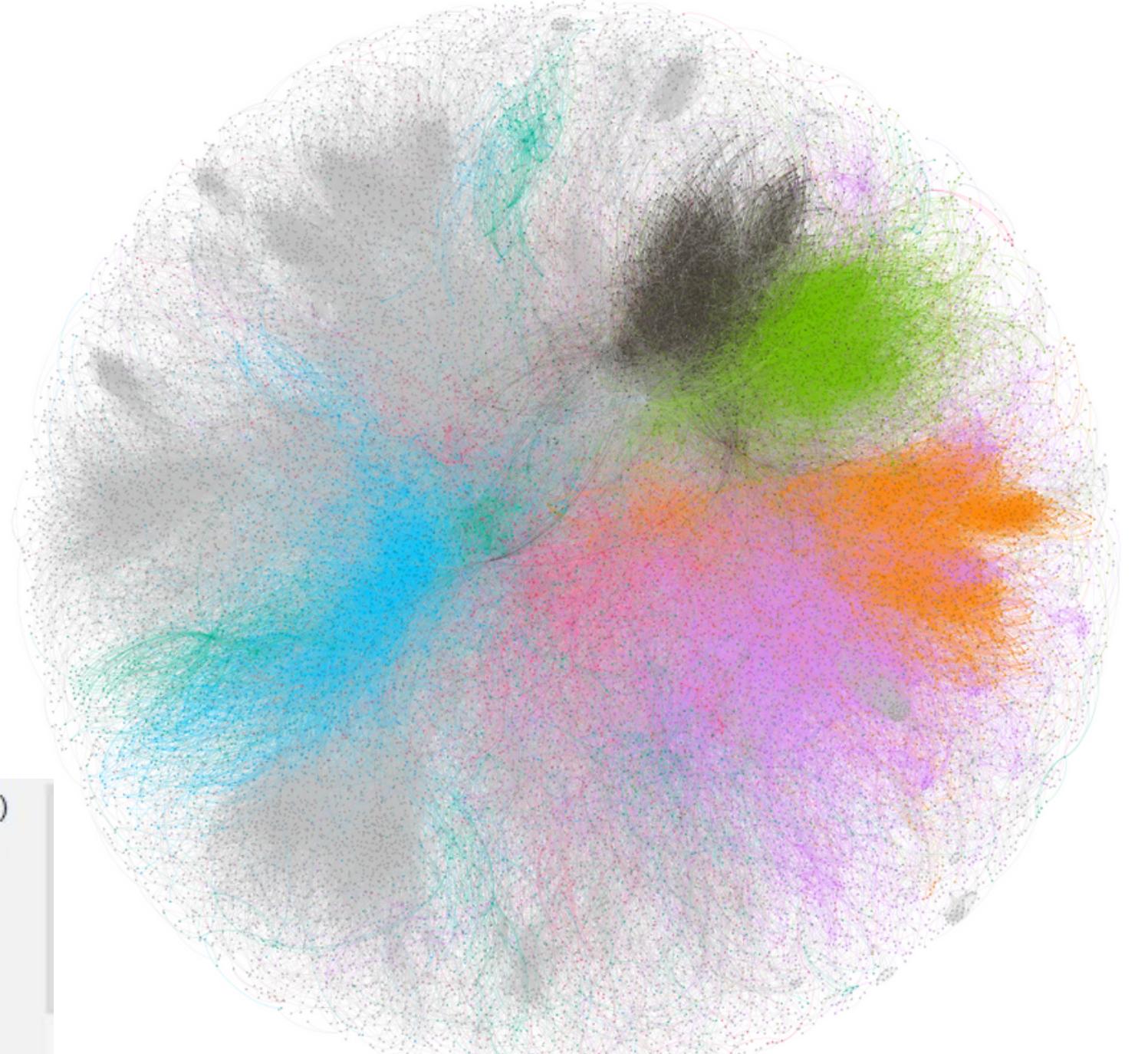
Louvain method

- 59 communities found
- 1 bigger community with more than 3000 nodes (15% of total nodes)
- 5 communities with more than 1000 nodes
- 11 communities with 500-1000 nodes

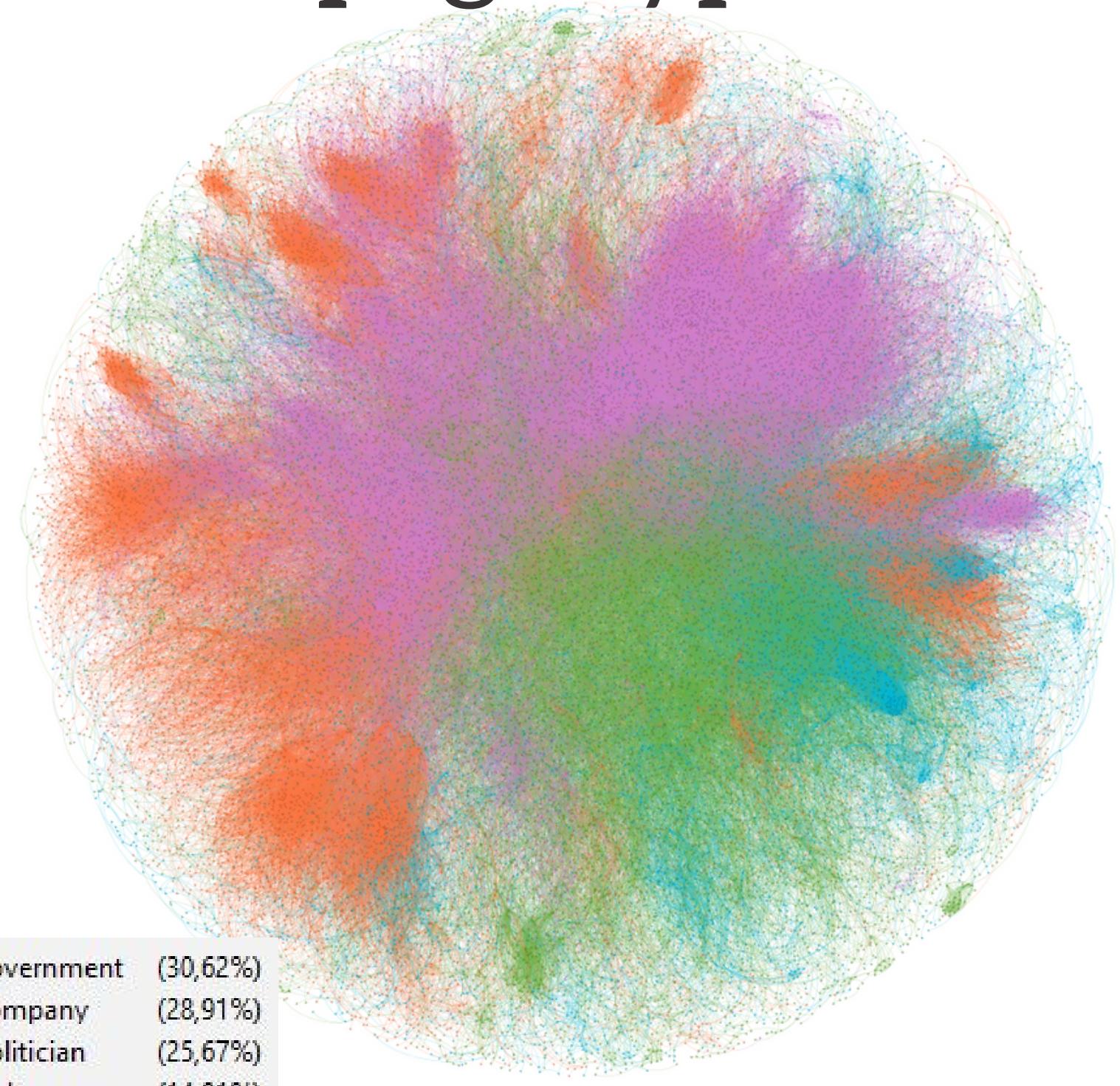


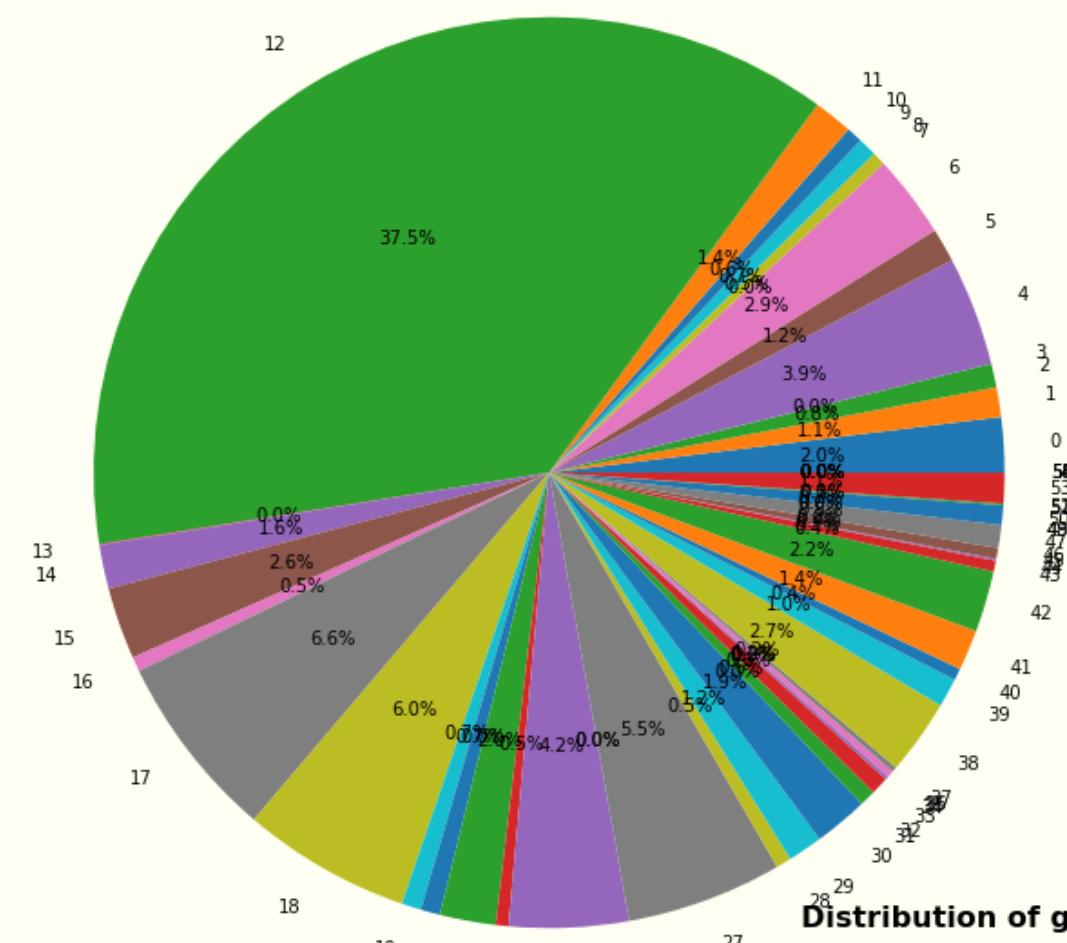
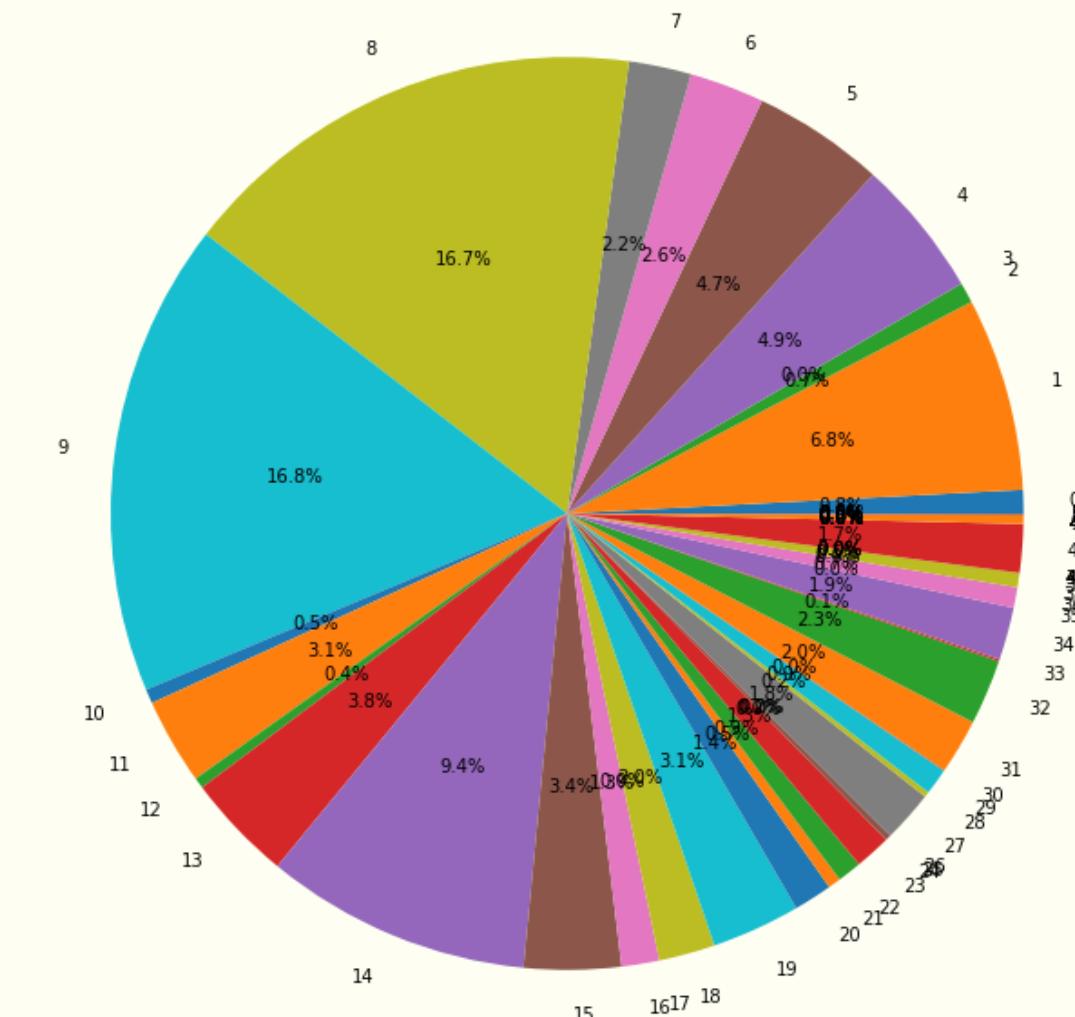
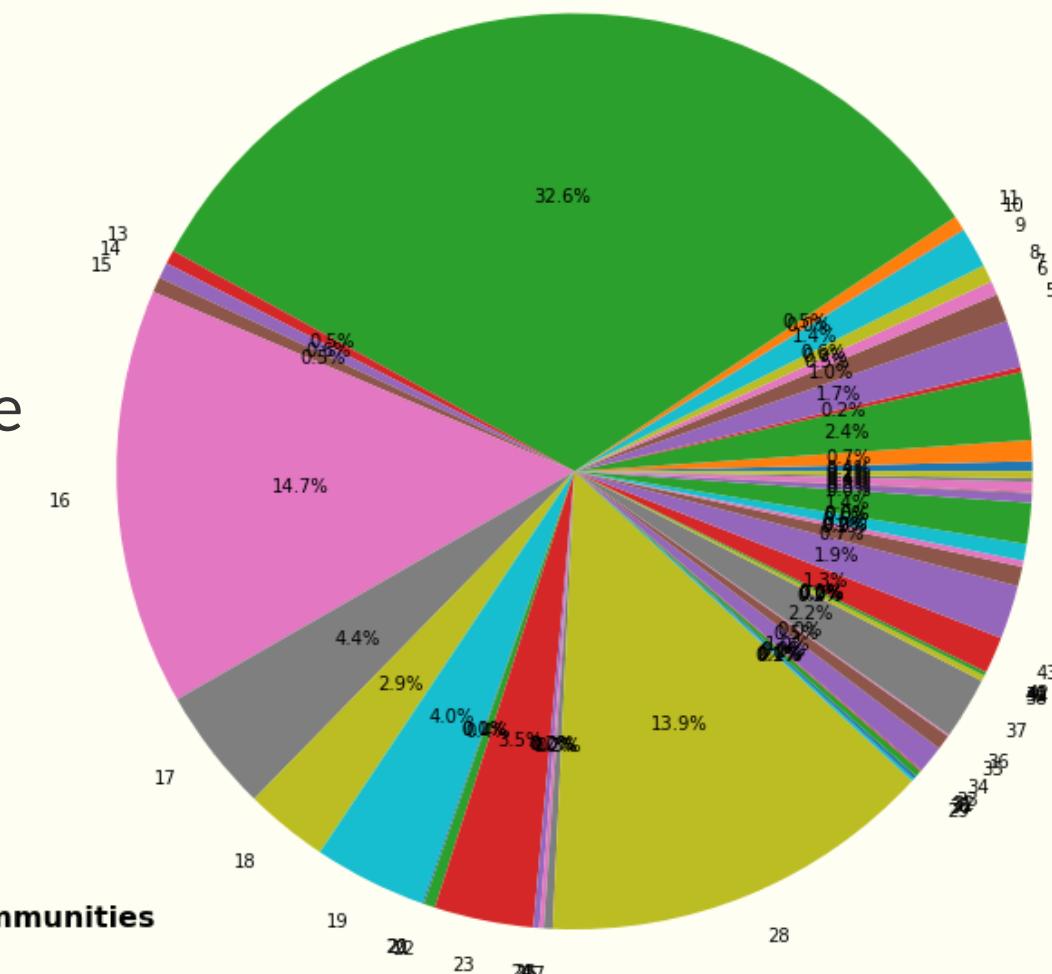
COMMUNITY DETECTION

communities

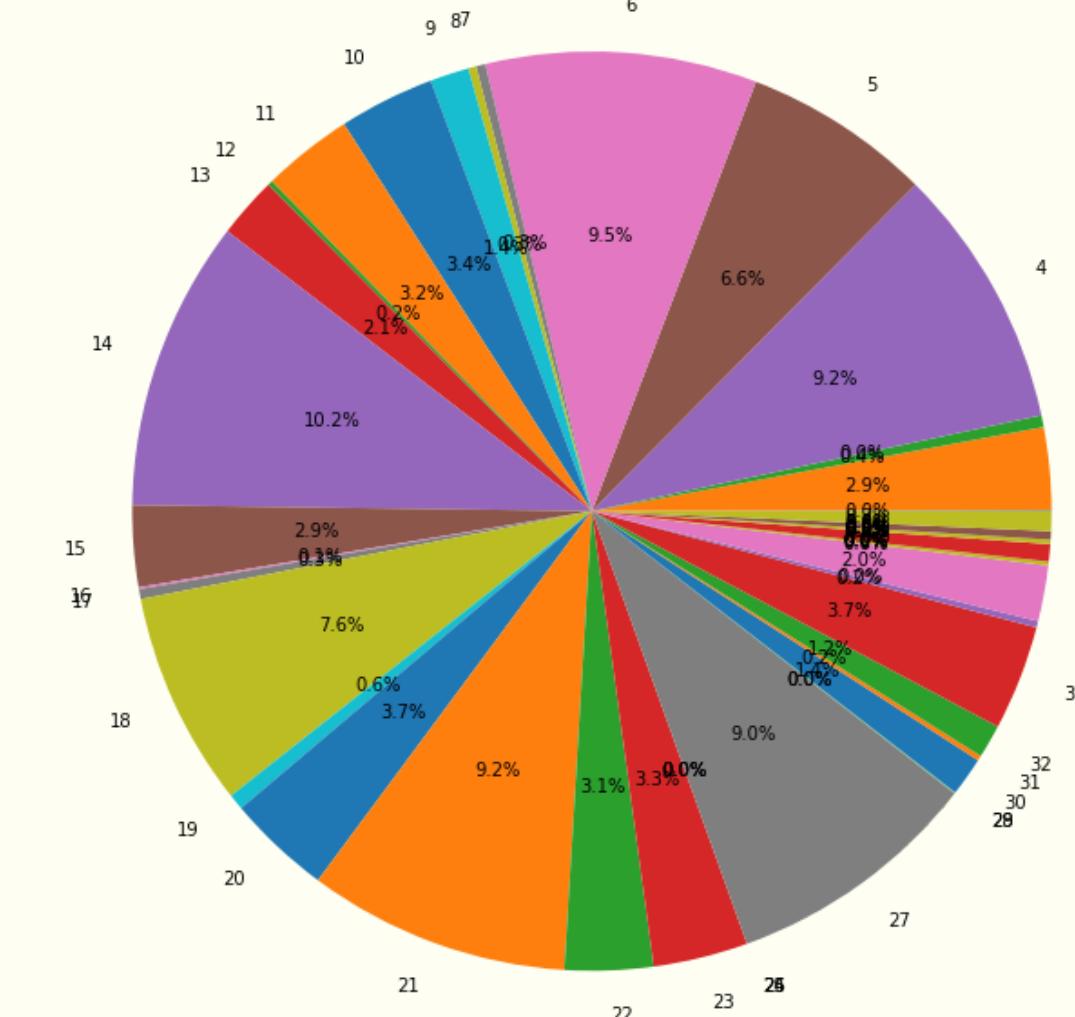


page type



Distribution of tvshow pages across communities**Distribution of government pages across communities****Distribution of company pages across communities**

A significant percentage of TV show pages and companies are in the same community (12):
 hypotheses: many TV shows are produced by a single broadcast television (like Mediaset in Italy, the same could be for NBC in USA) and there are companies financially supporting these shows in exchange for advertising.

Distribution of politician pages across communities

Walmart is a multinational retail corporation and one of the world's largest and most well-known retailers. Walmart is known for its wide range of products, including groceries, clothing, electronics, home goods and more.



"The Today Show" is a popular American morning television program that covers a wide range of topics, including current news, interviews, pop culture, health, lifestyle, and light entertainment.



Whole Food Market is an American supermarket chain that specializes in natural and organic foods. It is known for offering a wide selection of high-quality and often locally sourced produce, meat, seafood, dairy, and other grocery items.



"Home & Family" is a daytime television talk show that focuses on topics related to home, family and lifestyle. The show typically covers a wide range of subjects, including home improvement, cooking, crafts, health, parenting and celebrity interviews.

Further improvement:

looking at which users pages have in common expecting that pages with the same audience are likely to be connected.



THANK YOU FOR THE ATTENTION!