The Structure of a Decentralized Online Social Network: Mastodon

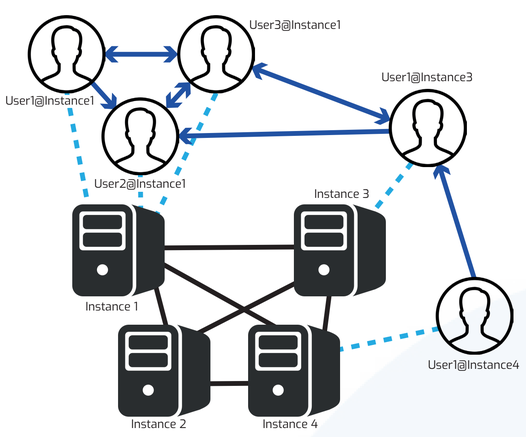
Matteo Zignani[1], Sabrina Gaito[1], Chantal Cherifi[2] and Gian Paolo Rossi[1]

[1]Department of Computer Science, University of Milan, Milan, Italy
{firstname.lastname}@unimi.it

[2] DISP Lab, University of Lyon 2, Lyon, France
chantal.bonnercherifi@univ-lyon2.fr

Motivations

We have witnessed the stunning growth of a plethora of web platforms based on the social relationships among their users. In becoming the most popular social networks, most have **drifted away from their original goals** and changed their business model as they faced problems like data monetization and data privacy. The utilization of user data for advertisement purposes is at the core of the public debate and impacts consumer trust in these online services. As a **reaction** to these concerns many tecno-activists and software developers have created **various forms of online social platforms that put social communication and user content at the heart of their actions**.



Mastodon: Decentralized Online Social Network

Mastodon is the newest and fastest-growing microblogging platform. Mastodon offers some interesting characteristics that make it eligible as a data source in different research fields:
a) unlike Twitter, Mastodon is **not centralized** but is made up of interconnected communities located in different servers, called *instances*;
b) **independently owned,** operated and moderated, each instance supports specific **interests and languages**;
c) users have a more detailed **control of the visibility** of their posts;
d) Mastodon **facilitates the access to information** related to the instances and the social communications among its members;
e) the bias given by the **recommendation systems is missing** in Mastodon. The only way to establish a connection is by searching an already known account through the search functions or by exploring the feeds of the instances in search of users with similar interests.

Dataset

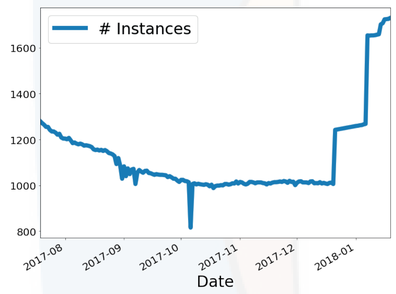
To get the dataset send an e-mail to matteo.zignani@unimi.it

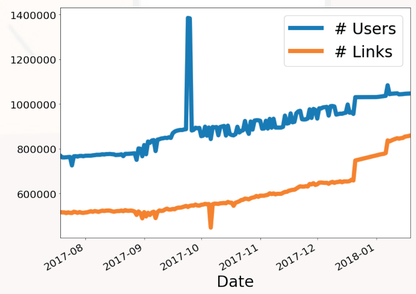
Structure and Evolution

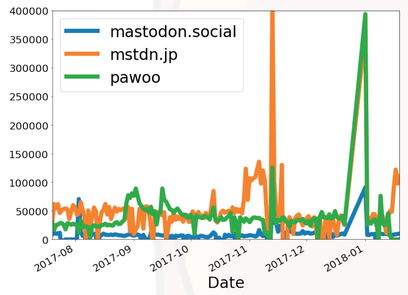
We crawled the structureof the social network resulting from the "follow" relationships among the Mastodon users.
We obtained a network made up of **479,425 nodes** and **5,649,762 directed links**. It covers 46% of users in Mastodon.

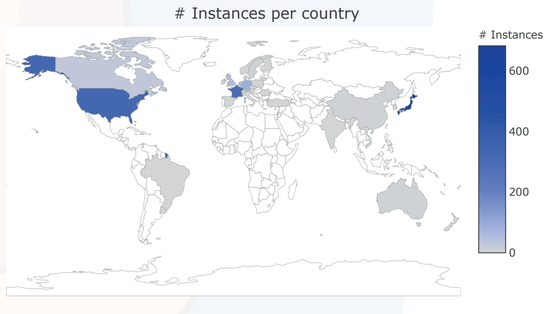
Instance name	Users (MD)	Users (N)	Coverage
pawoo.net	336182	131478	39%
mstdn.jp	156150	92322	59%
mastodon.social	124612	70409	57%
friends.nico	59918	17888	30%
mastodon.cloud	39803	31128	78%
mastodon.xyz	14207	6953	49%
octodon.social	9296	5080	55%
music.pawoo.net	8499	287	3%
mamot.fr	8269	5533	67%
social.tchncs.de	7896	2574	33%

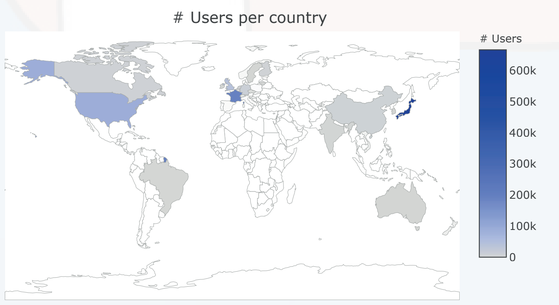
We were able to track the **evolution** of the network.



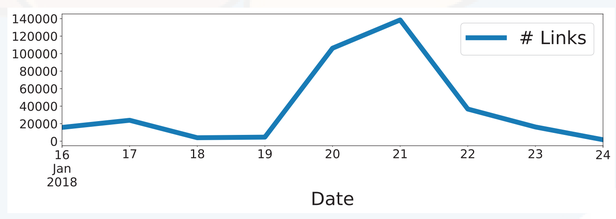








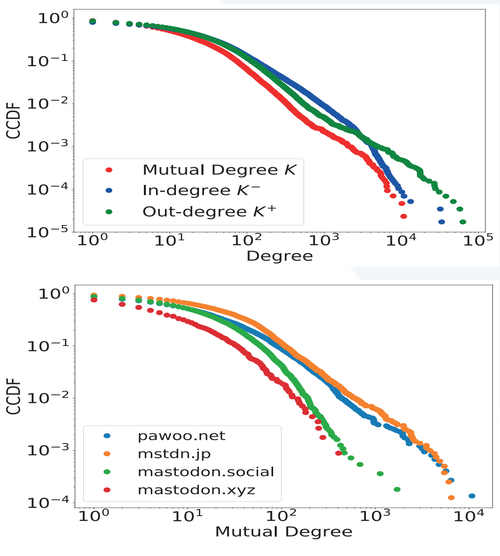




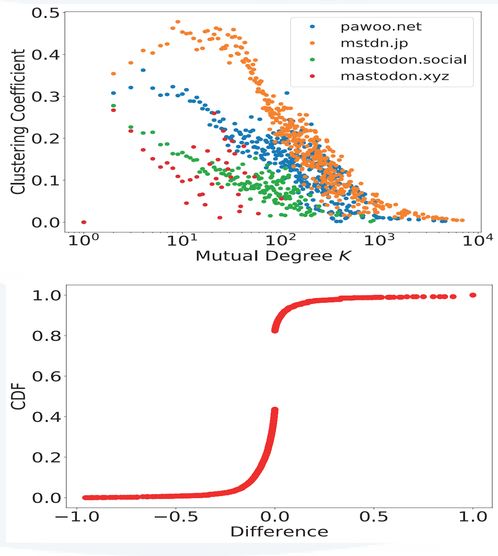
Instance Meta-data

For each instance, we obtained:
- Name: domain name of the instance.
- Users: the number of users registered on the instance.
- Connections: the number of connections between the instance and the other ones.
- Statuses: the number of posts published by users hosted by the instance.
- **Full description**: a description of the instance. It illustrates what are the topics and the prohibited contents and behaviors.
- **Topics**: a short list of topics the instance focuses on.
- **Geographical location** at a country-granularity

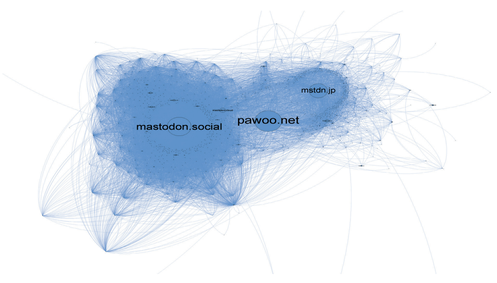
Findings



While the typical Twitter user follows more people than he/she has followed, in Mastodon users have a **more balanced behavior**. The centralized paradigm of Twitter makes the users' behavior uniform across country, while the decentralized approach makes it possible to build **subnetworks of people with different features**.



The Mastodon network shows a tightly clustered structure. The clustered structure of the network strongly depends on the instances, tendency of neighborhood's nodes of being clustered is limited within the instance boundary.



Despite the decentralized and fragmented architecture, Mastodon users keep connected to the core of the network and are able to search for friendships in other instances.