GROUP 07 - First Ideas

Data we will use

- We will use users as nodes
- For edges we have multiple options
    - Posts & replies
    - Following / Ignoring
    - Likes
    - We will probably use a combination of the ones above
- Timestamps will be important, since we want to identify temporal dynamics

How is a community defined?

- In terms of network measures / statistics
- From the Slides
    - Girvan-Newman Algorithm
    - Modularity Maximization (more efficient for large networks)
- Literature
    - Adaptive algorithms for detecting community structure in dynamic social networks [Nguyen et. al.] https://ieeexplore.ieee.org/abstract/document/5935045?casa_token=nsamlBAUm3wAAAAA:jUrVp9w8iKYrml_5X9oICvi5J8DtlPAtzp38Drmhs3aizkID90gin38vBsLPsY2-XOE3axj04A
    - Community structure in social and biological networks [Girvan et. al.] https://www.pnas.org/content/99/12/7821.short
    - Community detection algorithms: A comparative analysis [Lancichinetti et. al.] https://journals.aps.org/pre/abstract/10.1103/PhysRevE.80.056117

Split data into windows

- Granularity? Days, weeks, ...
- Look at how users get liked/disliked
- Look at how users comment on other users posts

Temporal Dynamics

- Compute network/graph statistics at different timings inside the dataframe and capture changes
- Community growth vs. community shrinkage
- Community with changing individuals vs. die-hard members
- How are the likes/dislikes of a post distributed over time / users

Further interesting questions

- Are users who have many follows/ignores more likely to get likes/dislikes in real time?
- Maybe a user with a lots of follows is no longer relevant so he gets very few likes / dislikes in relation to his follows (or less at then end of the month than at the start)
- https://hbr.org/2018/11/better-people-analytics : Influential people, check for static users and look for posts from them, maybe they are receiving more likes
- Echo-chamber effect: Do new posts receive likes etc mostly within the community or do they also reach other members?
- Are there any users that are famous in the community?
- Compare time dependent metrics with time independent state of network (= final state?)