

# Data Science and Big Data Analytics 16/17

## Exploring the Social Structure of the Zookeeper Project

Group 12

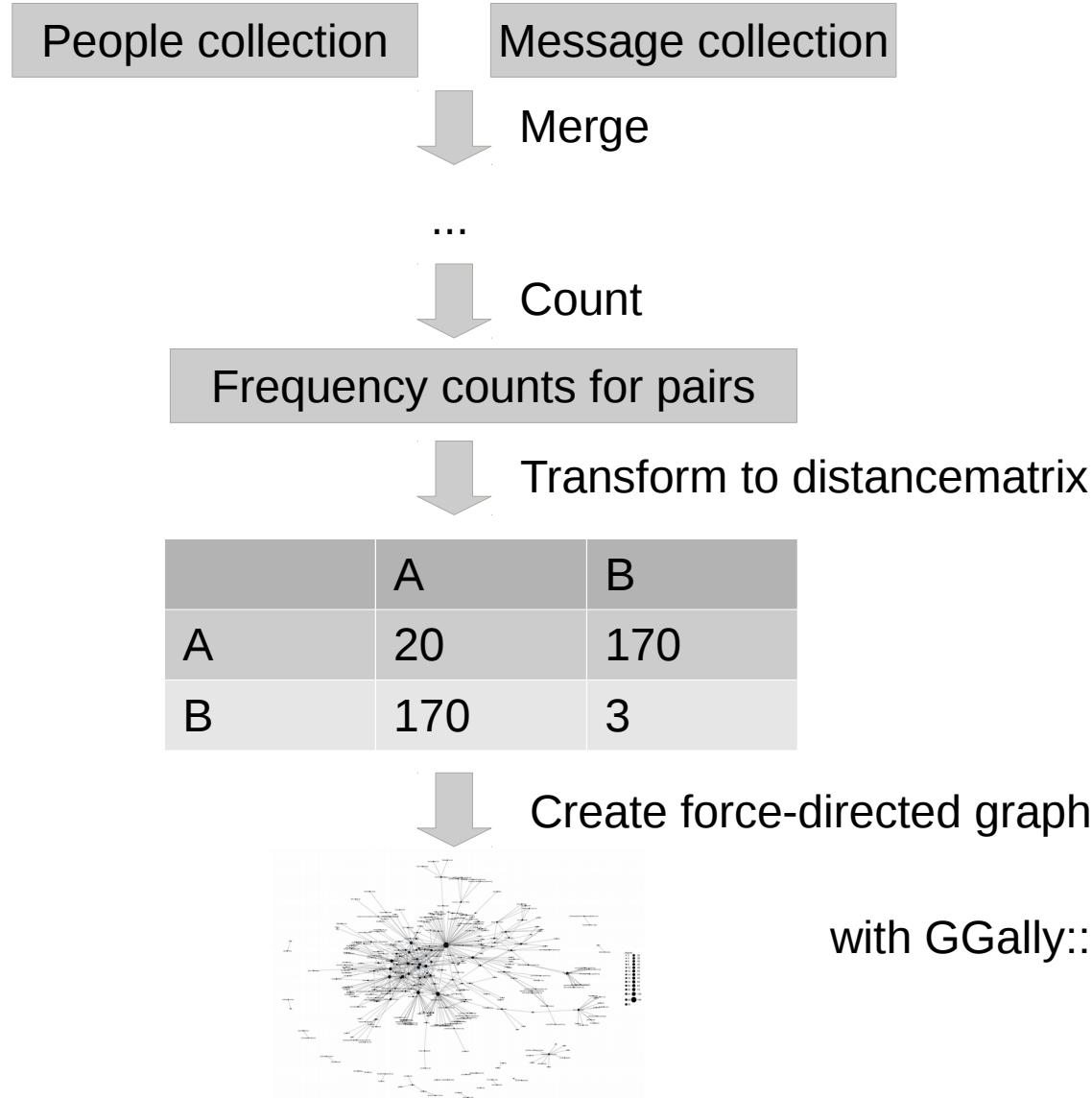
Lukas Steuernagel, Malte Sahrhage  
Markus Suhr, Andrew Lehr

31.01.2017

# Motivation

- SmartSHARK database: Contains information about the ZooKeeper software project.
- We set out to dissect the social structure of the ZooKeeper project.
  - Create networks with mails.
  - Examine who worked when and on which projects.

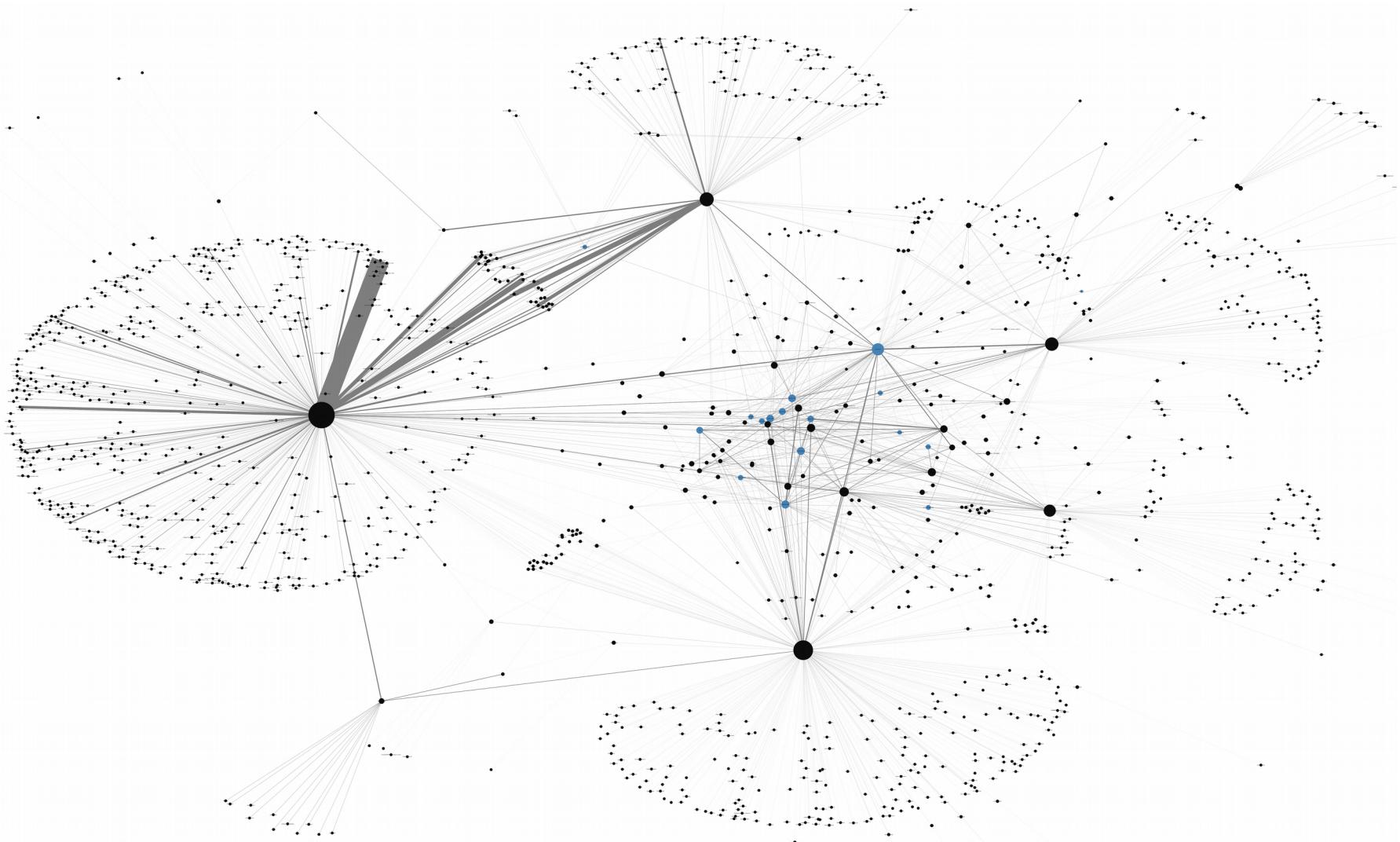
# Creating interaction networks



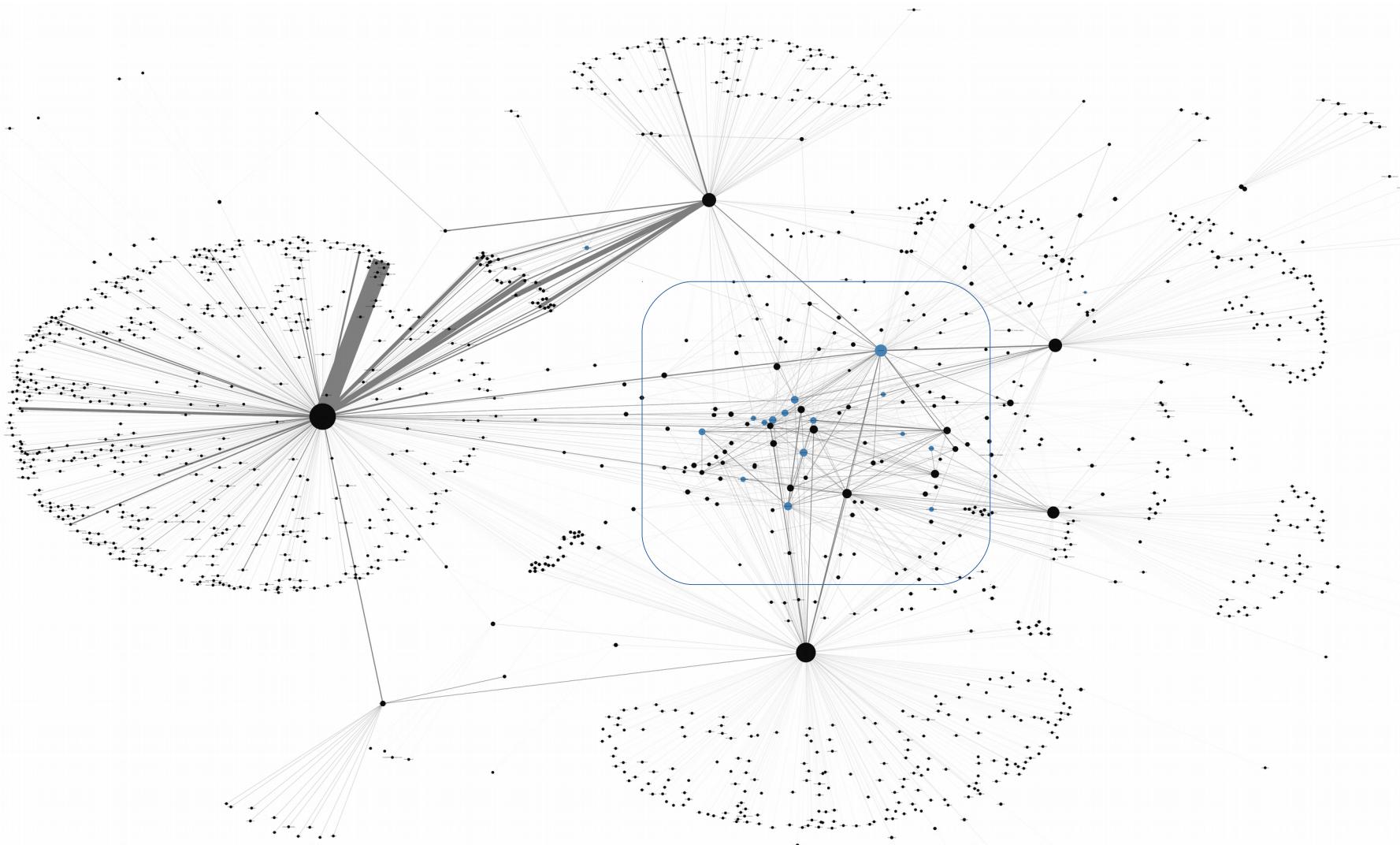
# Pairing bug-tracker issues with repository commits

- No direct key-relationship between collections
- Merging through issue ID mentioned in commit message
  - issue.external\_id = “ZOOKEEPER-123”
  - commit.message = “ZOOKEEPER-123. Changed something in this file”

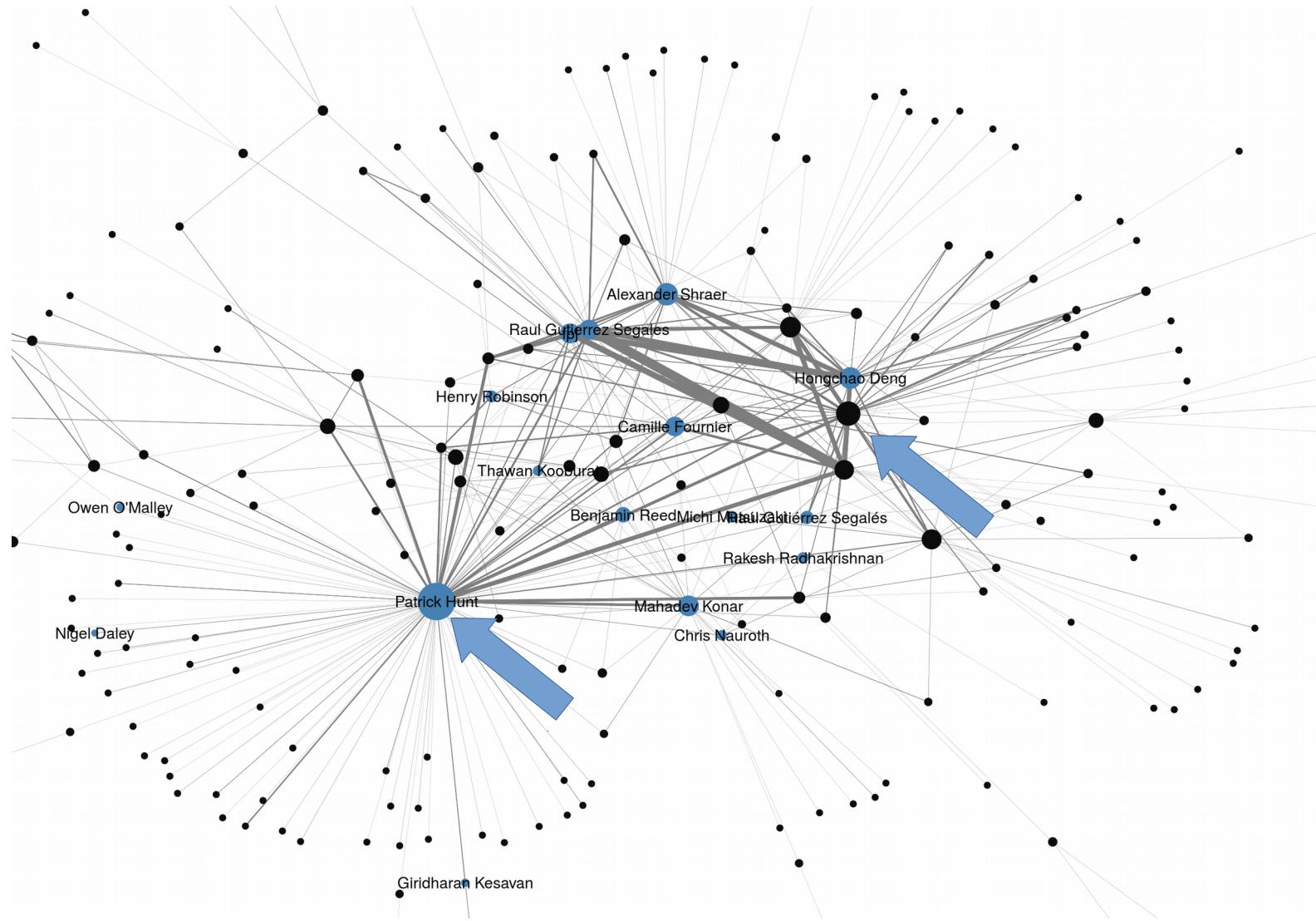
# The whole network



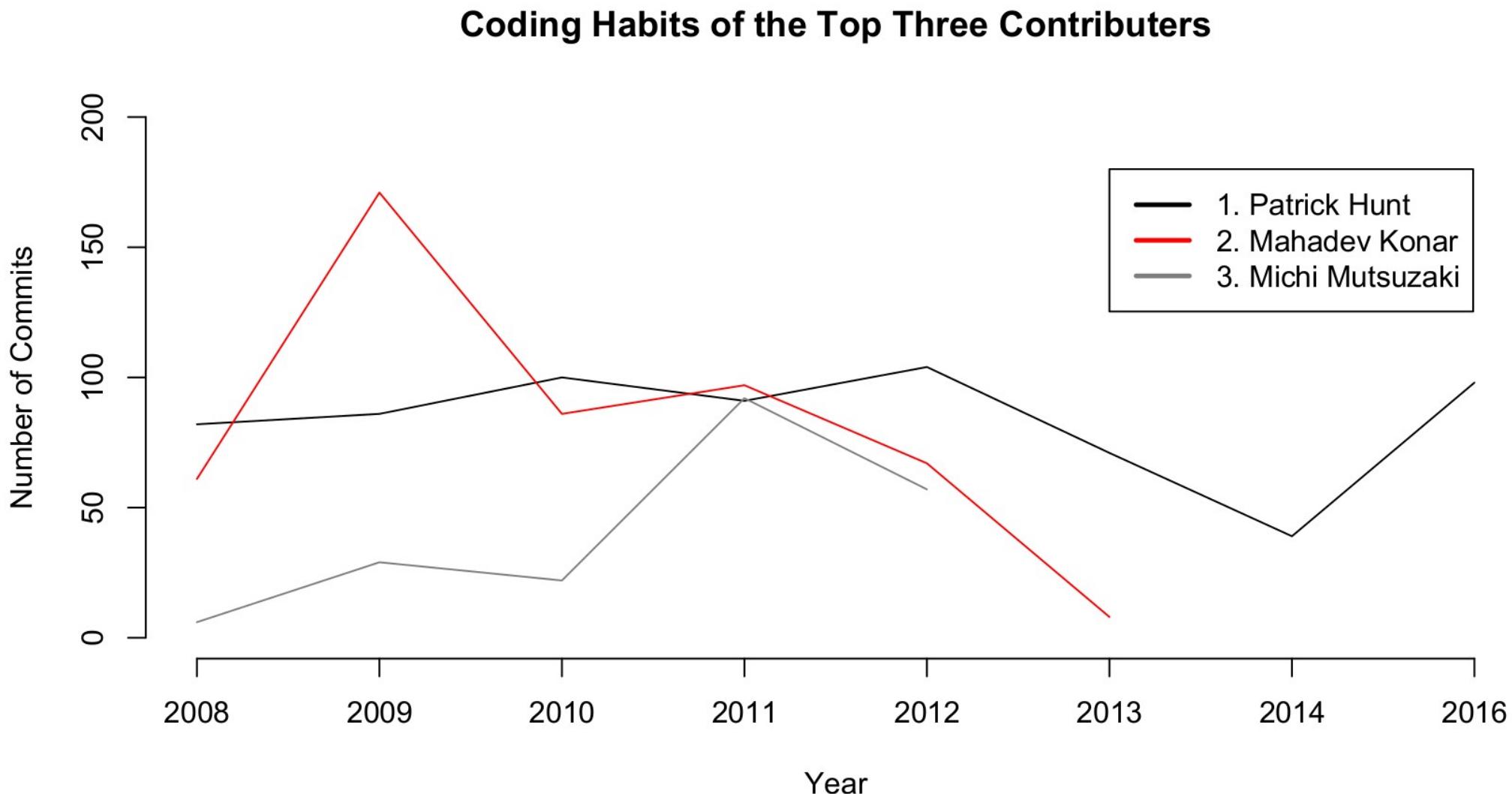
# The whole network



# The inner network

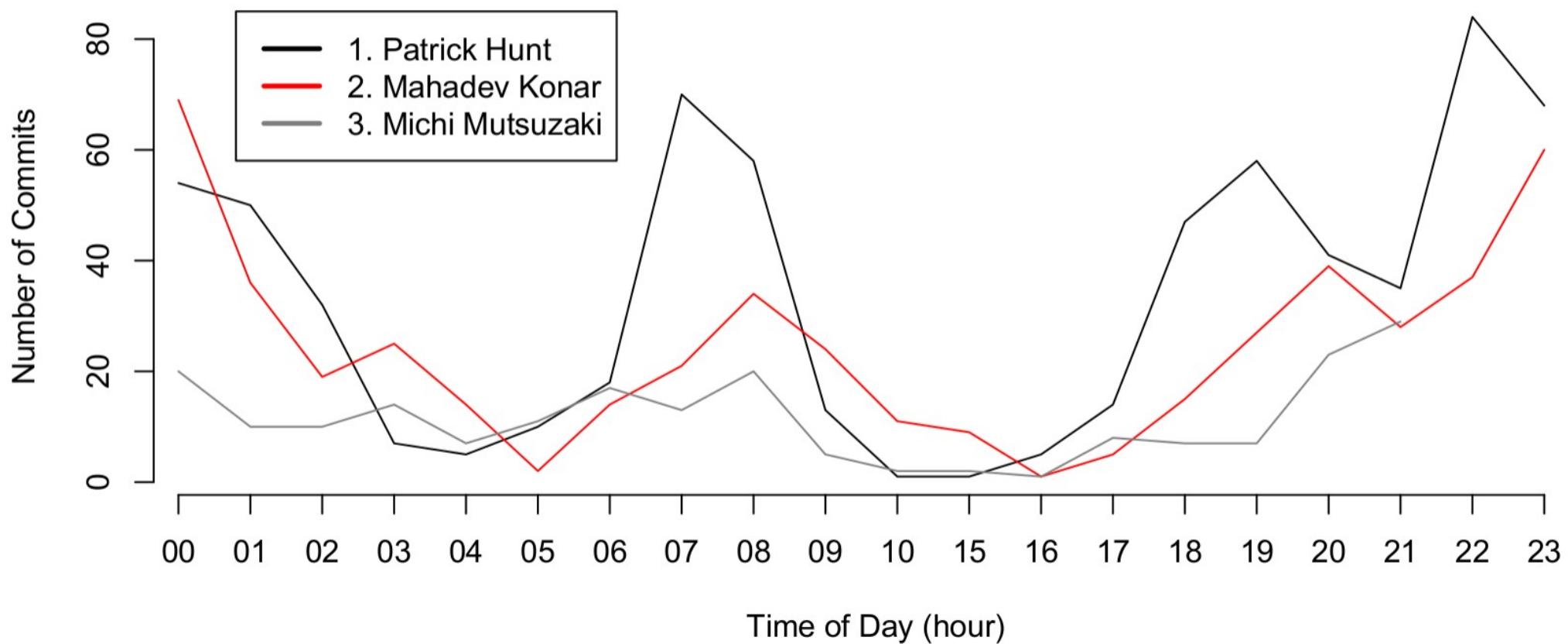


# Commit behaviour (1)



# Commit behaviour (2)

Coding Habits of the Top Three Contributors



# Questions?

