Process mining in IS MU

Jakub Hanko & Daniela Krúželová

File categories

- Scripts (3891)
- Modules (722)

Hypotheses to be verified

- Summer time vs. rest of the year
- .pm (modules) vs. .pl or .js (scripts)

Our typical activities in Git

- Big change (>200)
- Medium change
- Small change (<10)
- Rename
- Revert
- Quick fix

What were our steps?

- Discovering typical activities in our git repository
- Writing perl script to generate data into csv
- PM discovery by Disco

Our discoveries

- **Summer time:**
  - less activity in modules
  - average number of committed changes is higher in modules (new agendas)
  - more in Disco

Our discoveries

- **Scripts vs modules:**
  - In both cases small and medium commits were predominant
  - Ratio of small to medium commits in modules is almost 2 : 3 but in scripts it is 2 : 5
  - Quick fix and reverts usually don't happen after a big commit
  - Big commits were more often in modules
  - Less renames in modules

Our conclusions