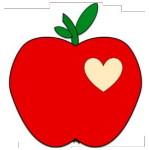
Using GitHub Analytics to Assess the Quality of Collaboration in Software Engineering Teams

Quan Le, Kiet Phan, Bowen Hui, Adara Putri



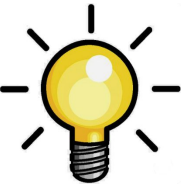

Computer Science
University of British Columbia, Kelowna, Canada

Motivation



Teaching goal:
Help students learn how to participate actively and communicate effectively in teams

- Team members have differing visions
- Individuals do not contribute equally to the work output



This work:
A framework for fairly assessing teamwork coupled with GitHub team analytics to detect collaboration issues

Related Literature

Team Process Models

Capstone Collaboration Assessments

Collaborative Work and Code Metrics

Weekly Assessments

Related Literature

Team Process Models

- Descriptive models about team processes
- Do not explain process transitions
- Do not explain how individual characteristics and behaviors influence team dynamics
- Relied on traditional data collection methods
- Field recommends use of digital traces

Capstone Collaboration Assessments

Collaborative Work and Code Metrics

Weekly Assessments

Related Literature

Team Process Models

- Descriptive models about team processes
- Do not explain process transitions
- Do not explain how individual characteristics and behaviors influence team dynamics
- Relied on traditional data collection methods
- Field recommends use of digital traces

Capstone Collaboration Assessments

- Most involve client projects
- Reported challenge of assessing individuals
- Lack assessment of team collaboration process
- Peer evals used as a proxy to assess team dynamics

Collaborative Work and Code Metrics

Weekly Assessments

Related Literature

Team Process Models

- Descriptive models about team processes
- Do not explain process transitions
- Do not explain how individual characteristics and behaviors influence team dynamics
- Relied on traditional data collection methods
- Field recommends use of digital traces

Capstone Collaboration Assessments

- Most involve client projects
- Reported challenge of assessing individuals
- Lack assessment of team collaboration process
- Peer evals used as a proxy to assess team dynamics

Collaborative Work and Code Metrics

- Advanced CSCW studies are limited to group interactions, not teams
- Code metrics focus on complexity
- Difficult to generalize across tech stacks and project types

Weekly Assessments

Related Literature

Team Process Models

- Descriptive models about team processes
- Do not explain process transitions
- Do not explain how individual characteristics and behaviors influence team dynamics
- Relied on traditional data collection methods
- Field recommends use of digital traces

Capstone Collaboration Assessments

- Most involve client projects
- Reported challenge of assessing individuals
- Lack assessment of team collaboration process
- Peer evals used as a proxy to assess team dynamics

Collaborative Work and Code Metrics

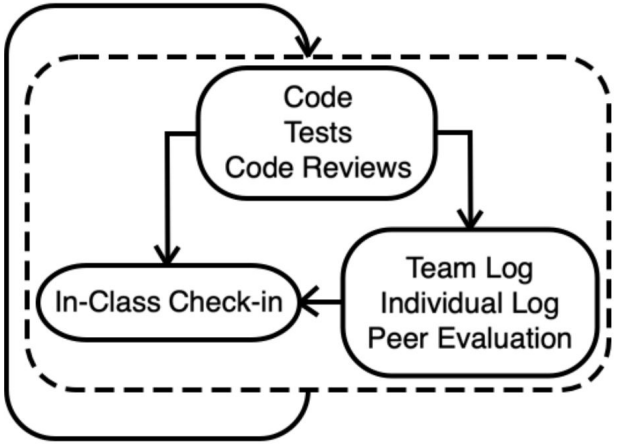
- Advanced CSCW studies are limited to group interactions, not teams
- Code metrics focus on complexity
- Difficult to generalize across tech stacks and project types

Weekly Assessments

- Some researchers use weekly assessments to obtain longitudinal data
- Such evals must be short and easy to complete
- Example measures:
  - Workload contributions
  - Belongingness
  - Team functioning

Proposal: Assessment Triangulation Framework

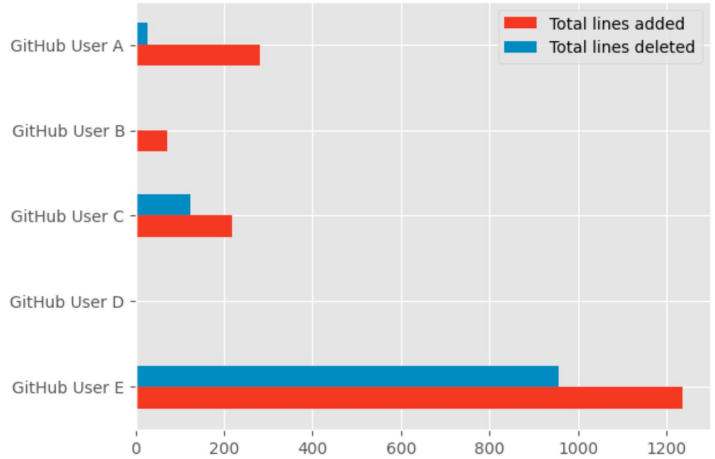
- **Observed data:**
  - Combines repository activity and collaboration analytics into a **PR report**
- **Self-reported data:**
  - Team logs
  - Individual logs
  - Peer evals
- **Meetings:**
  - Weekly in-class checkin's
  - Resolves discrepancies
  - Discusses progress and plans



Overview of PR Activities

- Broad overview of productivity at the individual and team levels
- Insights about the team's code development processes
- Quick comparison of contributions relative to teammates

Contributor	PRs (Merged / Not Merged)	Commits	Lines added	Lines deleted	Lines contributed	Files changed
GitHub User A	2 (2 / 0)	6 (3.0/PR)	280 (140.0/PR)	26 (13.0/PR)	254 (127.0/PR)	11 (5.5/PR)
GitHub User B	2 (2 / 0)	4 (2.0/PR)	72 (36.0/PR)	1 (0.5/PR)	71 (35.5/PR)	4 (2.0/PR)
GitHub User C	3 (0 / 3)	8 (2.7/PR)	217 (72.3/PR)	125 (41.7/PR)	92 (30.7/PR)	8 (2.7/PR)
GitHub User D	0 (0 / 0)	1 (0/PR)	0 (0/PR)	0 (0/PR)	0 (0/PR)	0 (0/PR)
GitHub User E	2 (0 / 2)	32 (16.0/PR)	1,236 (618.0/PR)	957 (478.5/PR)	279 (139.5/PR)	5 (2.5/PR)
Total	**9 (4 / 5)**	**51 (5.7/PR)**	**1,805 (200.6/PR)**	**1,109 (123.2/PR)**	**696 (77.3/PR)**	**28 (3.1/PR)**



PR Details

- Insights on coding and development practices
- Clarifications on timeline of features and updates

PR Title: Browse screen reserve

PR Author: GitHub User B

Description: Removed the reserved post from browse screen and book mark screen

Status: Open
Number of commits: 6
Lines added: 13
Lines deleted: 6
Lines contributed: 7
Files changed: 2
Reviewers: GitHub User C , GitHub User D , GitHub User B
Created at: 2024-02-18T17:50:52Z
Closed at: Still Open

Commit History
@GitHub User B - Fix bug
@GitHub User B - Remove whitelist after unsubmit
@GitHub User B - Add tests
@GitHub User B - resolve comment
@GitHub User B - Resolve comment

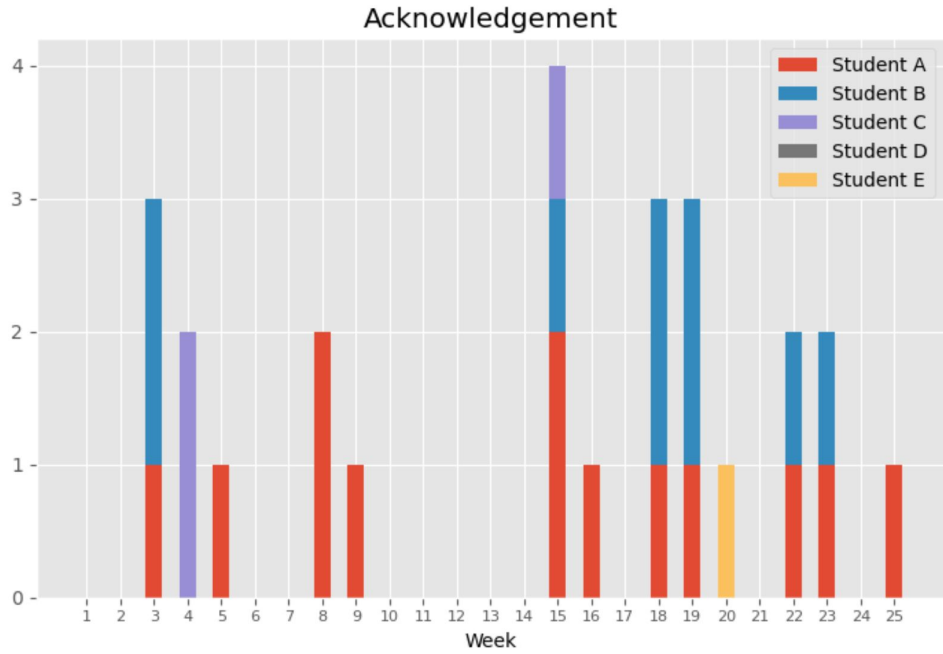
PR Review Comments

- Insights on code review participation and contributions
- Feedback provision practices
- *Future work: comment extraction and analysis*

Contributor	Comments	Review Replies	Words per Comment	Reviews
GitHub User A	2	0	9.0	4
GitHub User B	0	0	0	1
GitHub User C	1	0	12.0	1
GitHub User D	2	0	16.5	3
GitHub User E	2	0	12.5	2
Total	**7**	**0**	**12.6**	**11**

PR Review Interactions

- Insights on "invisible" work
  - Pair programming
  - Assistance or conflict notes
  - Project management
- *Future work: bipartite graph*

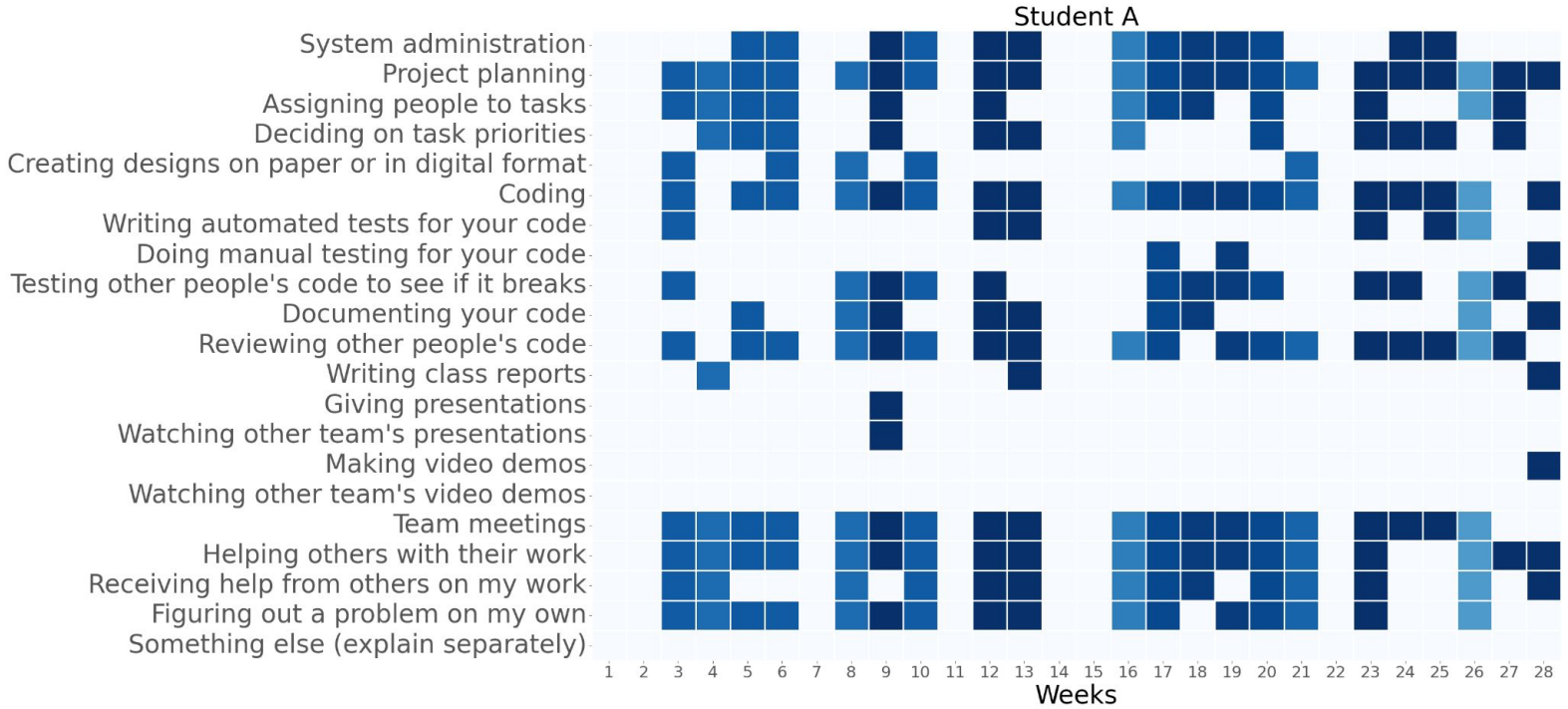


Distributions of Individual Contributions

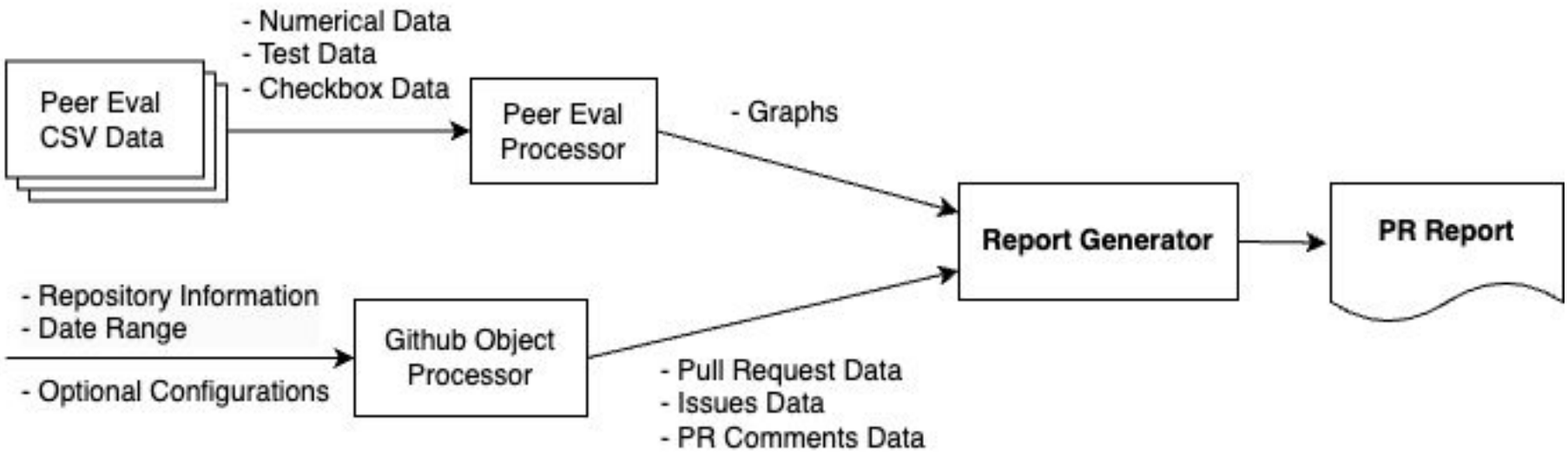
- Average % of workload contributions
- Also available:
  - Talking time
  - Decision making
- Shows the dominance relationships within a team

Self-Assigned Tasks Completed

- Work output and distribution consistency



System Architecture



Research Questions

1. What are the administrative gains afforded by the use of PR reports?

2. What are the potential risks of using these PR reports as part of the assessment process?

3. What information should be used in place or in addition to the analytics in the PR reports?

Course Context

- Fourth-year undergraduate Software Engineering Capstone course
  - Two semesters between September and April
  - 100+ Computer Science students formed 20+ teams
    - 1 instructor
    - Limited TA support
- Three course evaluation components:
  - Team component
  - Individual component
  - Client component

Pilot Study: TA Experience

- 4 Teaching Assistants as primary participants
- Qualtrics survey to the TAs to provide anonymous feedback about the PR reports
  - 10 structured questions (Yes/No, even-point Likert)
    - Summarized numerically
  - 10 open-ended questions (explanations and suggestions for improvements)
    - Thematic analysis

Results

Pros

- Insightful
- Supplemental
- Efficient
- Accessable

PR Reports gave ***accessible*** and ***insightful*** information for grading which proved ***supplemental*** to assessing students making the process more ***efficient***

Cons

- Misleading
- Unclear

PR Reports were sometimes found to be ***misleading***, with some inconsistencies and inaccuracies making the reports ***unclear***

Discussion and Future Work

- New features desired:
  - Improved details and data filtering
  - Detecting tests and computing test coverage
  - GitHub project activities such as issue creation and assignments
- Limitations
  - Reports might become overly excessive with added information