

Using CVS Historical Information to Understand How Students Develop Software

Y. Liu, E. Stroulia, **K. Wong**

D. German

Department of Computing Science
University of Alberta

Department of Computer Science
University of Victoria



JRefleX Project

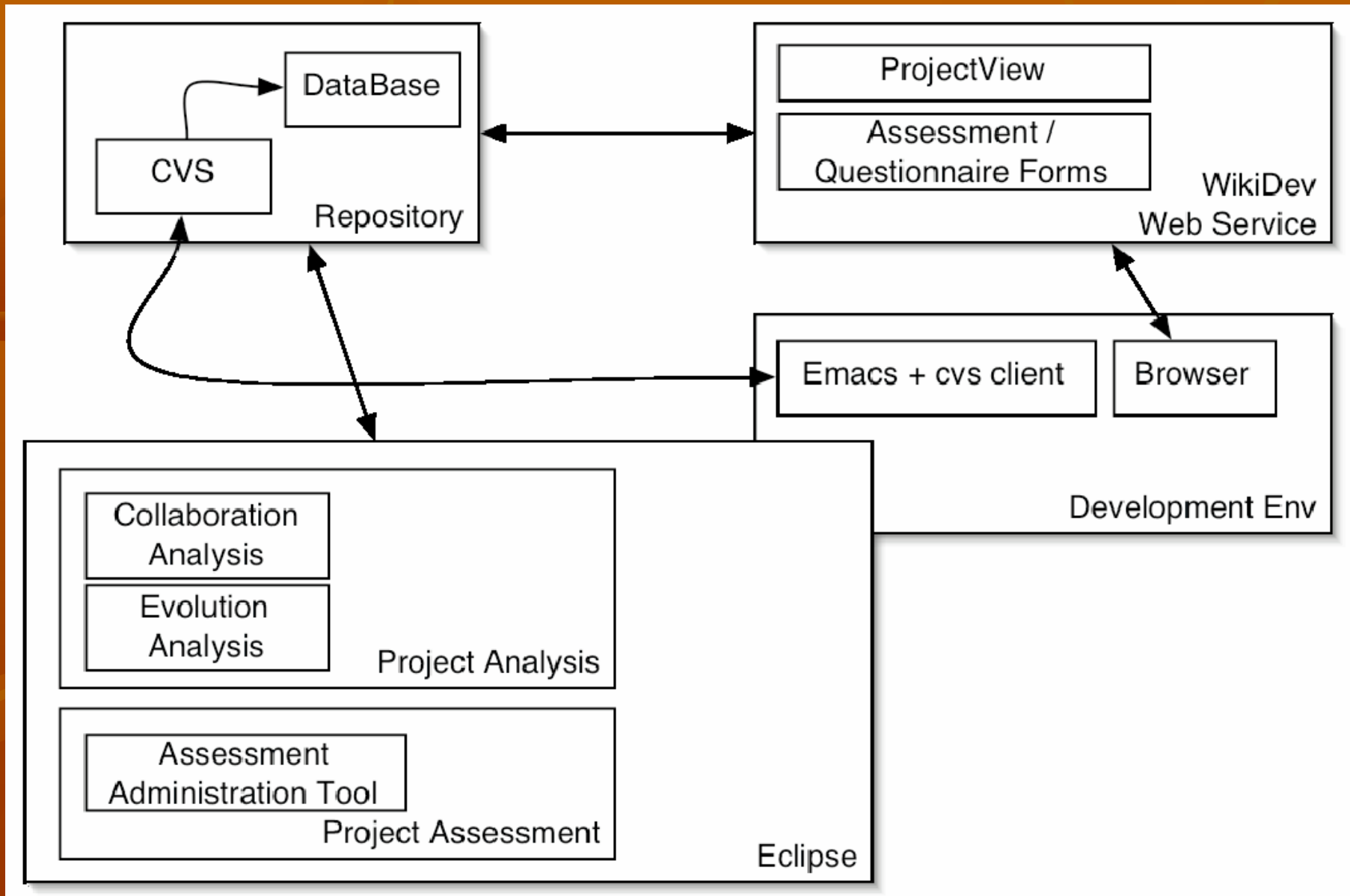
2004-05-25

Goals

- Enable in-depth understanding of team projects:
 - team collaboration patterns
 - individual workload and work pattern
 - code module evolution trends
- Assist instructors/managers:
 - monitoring teams
 - providing relevant/timely advice

Constraints

- Tightly integrated with development
 - not an extra effort, unobtrusive
- Current information
 - “real-time”, as the project is developed
- Multi-perspective
 - include information from revision control, work products, team perceptions, instructor assessments



301wiki - Home Page

http://localhost/~stroulia/wikidev/index.php

RecentChanges FindPage add page: LikePages BackLinks

Home Page

Welcome to the CMPUT 301 Moo.

Schedule 8 DemoTimeForProjectInterfacePrototypes
 *NOTE!! This is the schedule for Project part III Demo

NewDemoTime?

All are welcome. Sign in is required before edit.

Don't forget to sign the [Consent Form](#) permitting Drs. Stroulia and Wong to use your anonymous CVS and Wiki usage data for research purposes! Thank you!

[OldHomePage](#)

- Note: Plugins are disabled in this wiki so some features will be unavailable

Last edited on Wednesday, March 26, 2003 4:22:30 pm.

Edit PageHistory Diff Info Sign In

Top page

c301g62 Wiki! - Home Page

http://localhost/~stroulia/wikidev/c301g62/index.php

RecentChanges FindPage add page: LikePages BackLinks

Home Page

!!!Guys please check your page to see if there is stuff you need to do

Welcome To the MORcal Development Wiki, brought to you by MicroSpon

- [wno's using what](#)
- [Group Notes](#)
- [Documentation](#)
- [Wmvc Framework](#)
- [Default Crap](#)
- [Extended Apps](#)
- [Extended Models](#)
- [Extended Views](#)
- Our website <http://www.telusplanet.net/~tonykcm/c301/>

User Pages

- [Blaine's Page](#)
- [Den's Page](#)
- [James's Page](#)
- [Yury's Page](#)

Project home page

c301g62 Wiki! - Main View

http://eleni-stroulias-computer.local/~stroulia/wikidev/:301

RecentChanges FindPage add page: LikePages BackLinks

Main View

Viewing Current Bugs

Current Bugs

Assigned To: Nobody Yet

Bug Addressed? May 7, 2003, 11:07 am

Bug Type: Logic/Function

Priority: Low

Comments

Add Bug

Checking out MainView Revision: 1.20

```
import java.util.*; import java.awt.*; import javax.swing.*; import java.awt.event.*;
public class MainView extends WmvcView{
//private Month tempmonth; private CalendarModel tempcal; //private CalendarPage calpage;
/*Constructor*/
```

Source Control

Get Revision: 1.20

Autogenerated Wiki page of a CVS file

c301g62 Wiki! - groupnotes

http://eleni-stroulias-computer.local/~stroulia/wikidev/c301g62

RecentChanges FindPage add page: LikePages BackLinks

groupnotes

Group Notes

THIS PAGE IS SET UP FOR NOTES THAT YOU WANT THE WHOLE GROUP TO SEE

still to do

- parser
- undo/redo
- file export
- some kind of storage class
- move users to JCheckBoxMenuItem ??
- todos?
- two level help, with screenshots

additional

- highlight days on the year view / make them clickable to the day view

Edited Wiki page

Collaboration Analysis

- Understanding team collaboration:
 - What is the “collaboration profile” of each **team**?
 - What works, what doesn't?
 - How does each **developer** contribute?
 - What files does the developer affect?
 - With what kinds and numbers of operations?
 - How does each **work product** evolve?
 - How does it change?
 - Which and how many developers affect it?

Case Study

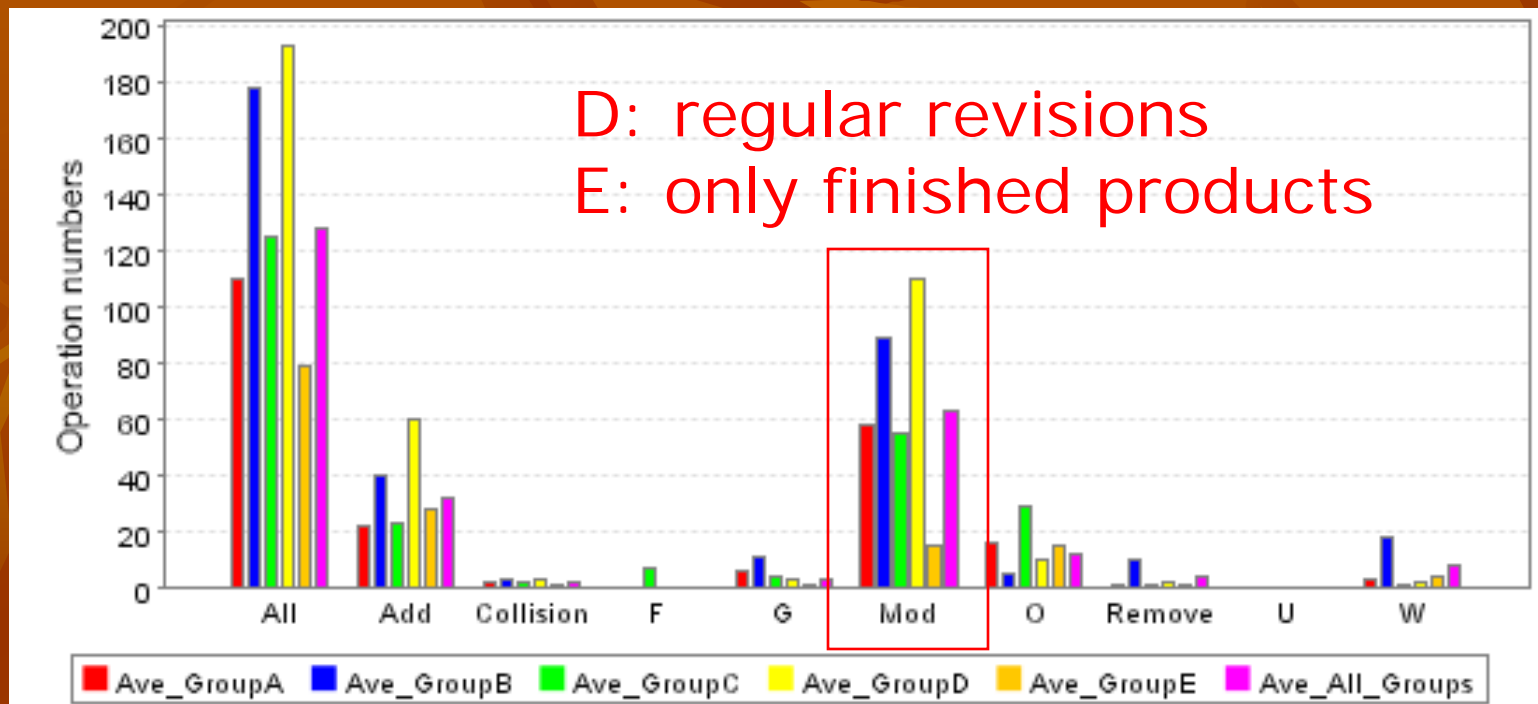
- Small undergraduate student teams:
 - 3 to 4 developers each
- Development process:
 - common project, three delivery stages
- CVS environment:
 - records evolving work during process
- Novices:
 - new to using CVS in teams
 - little project management experience

Comparing Across Teams

- Questions:
 - What is the distribution of CVS operations?
 - How fast do they start?
 - How long is their development process?
 - How many idle days?
 - ...

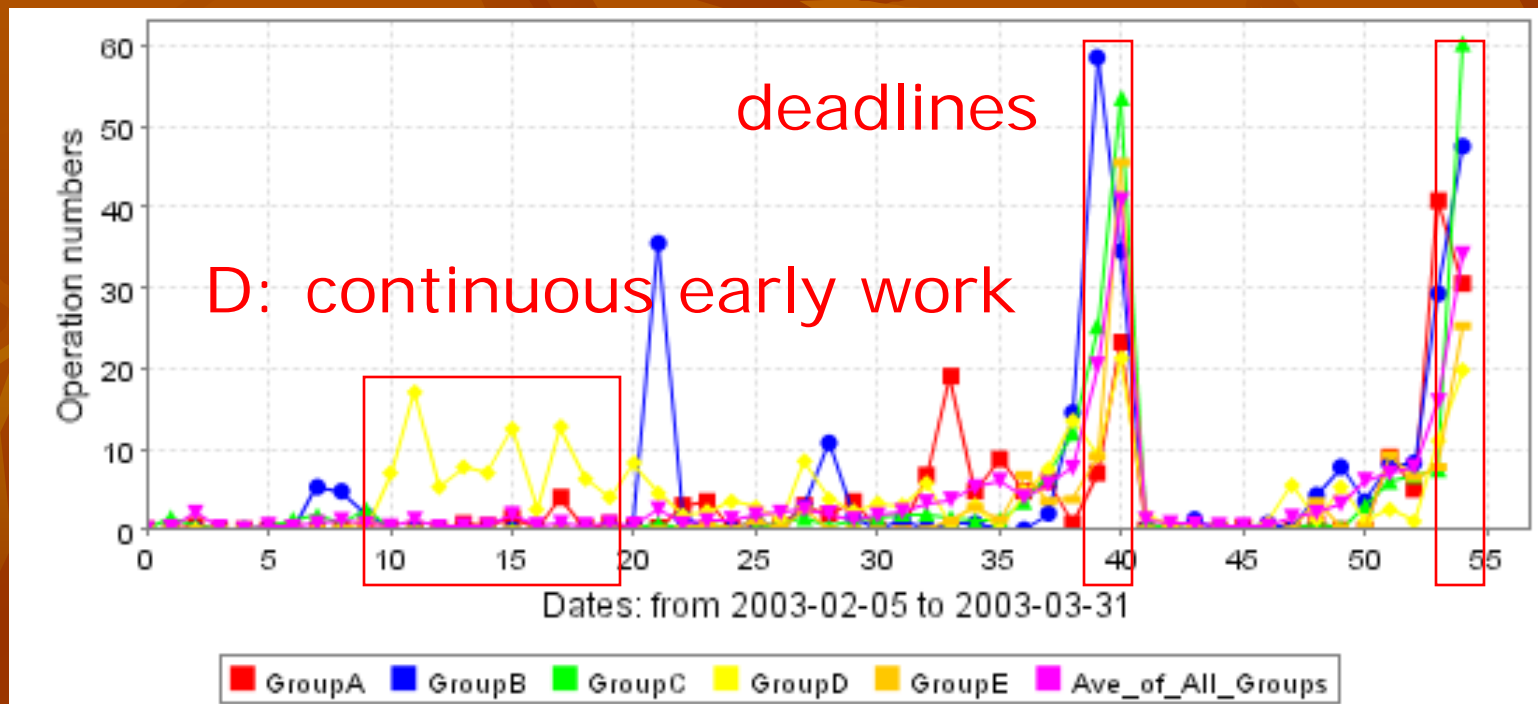
Comparing Across Teams

- Average numbers of operations of all types, for the teams:



Comparing Across Teams

- Total numbers of operations over time, for all teams:

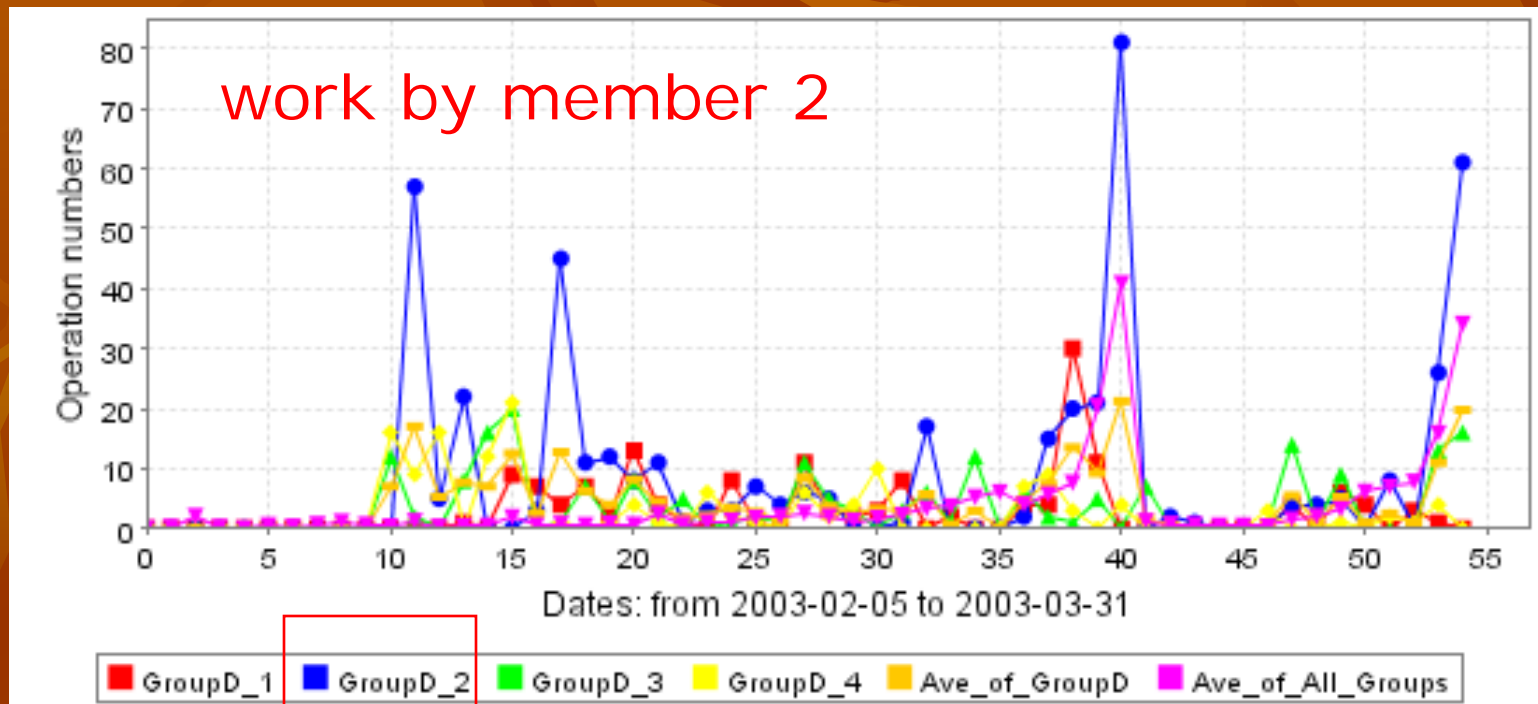


Members within Team

- Questions:
 - Who contributes?
 - In what role(s)?
 - stub writer, tester, debugger
 - ...

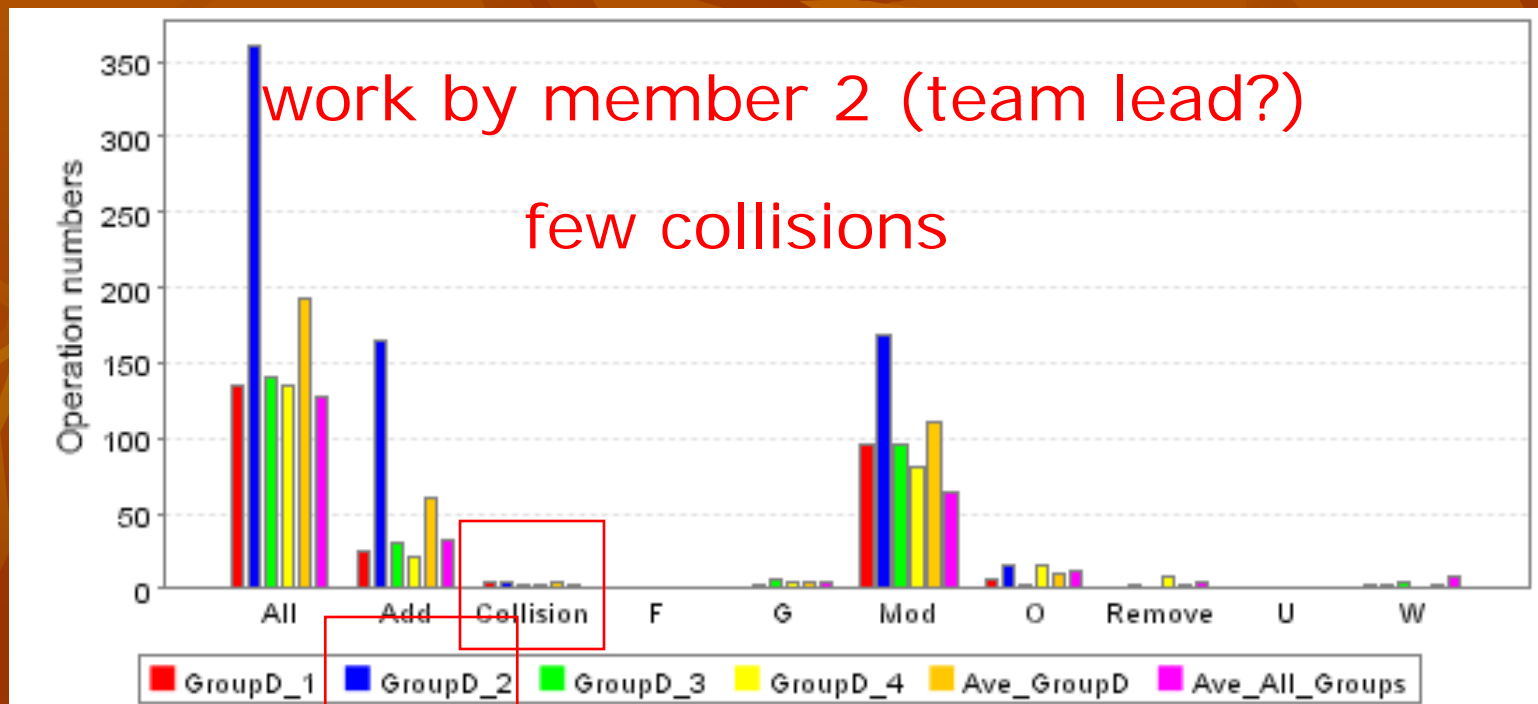
Members within Team D

- Total numbers of operations over time, for members of team D:



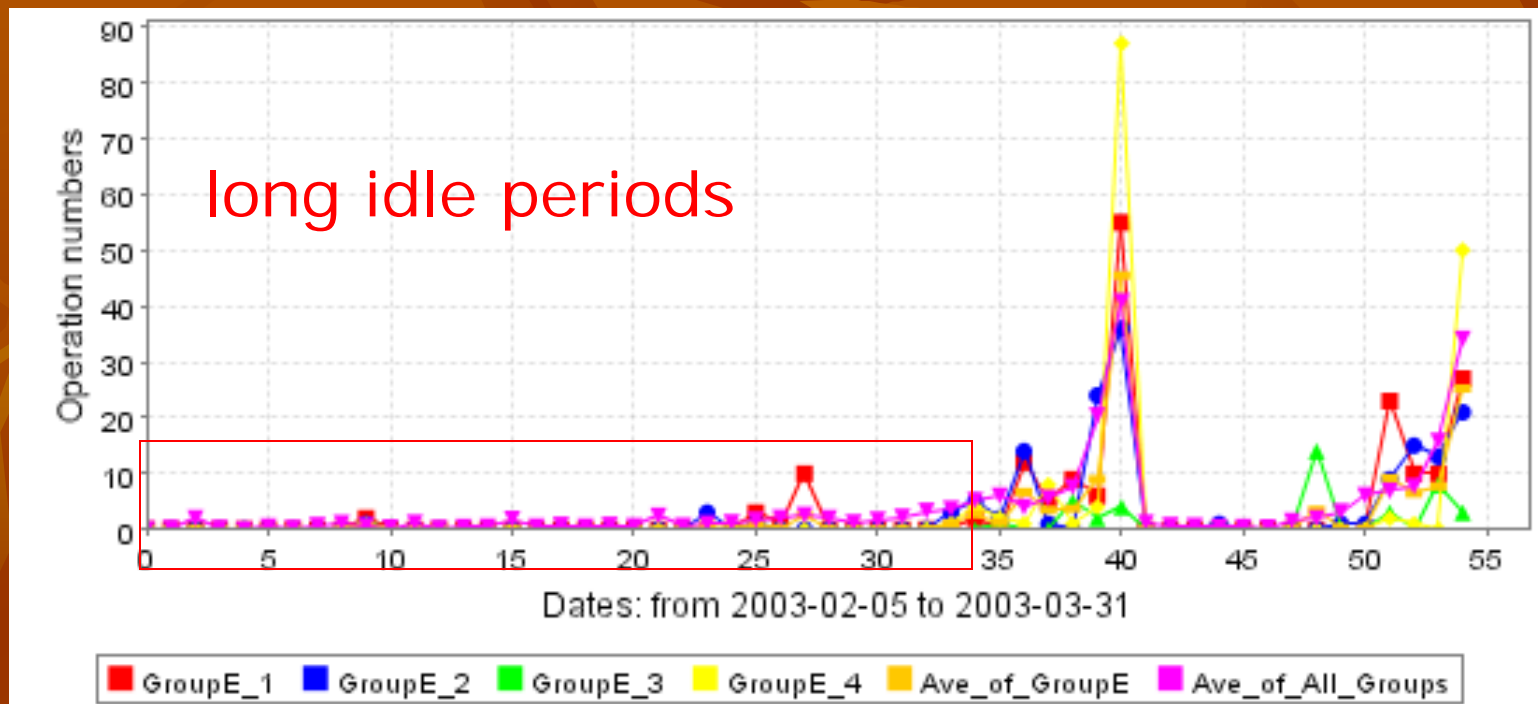
Members within Team D

- Numbers of operations of all types, for members of team D:



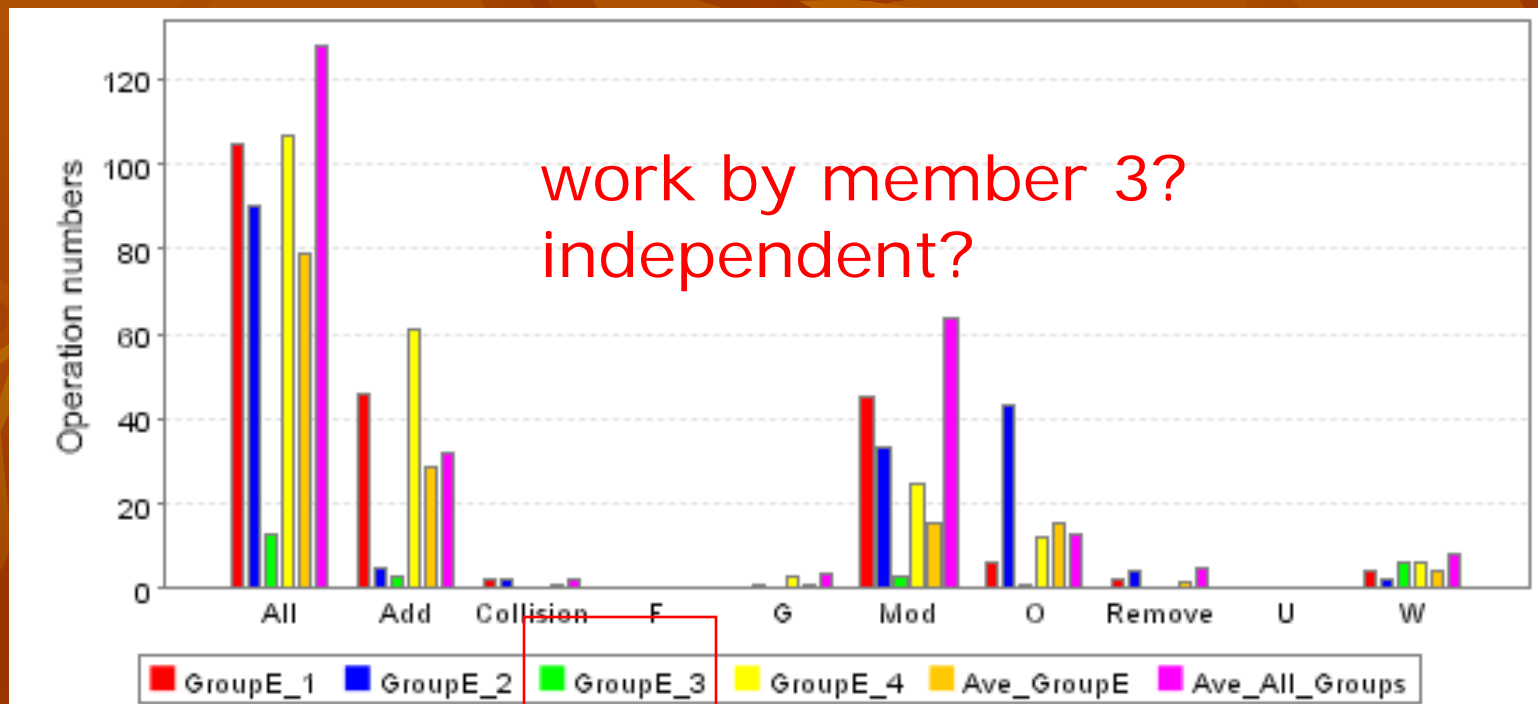
Members within Team E

- Total numbers of operations over time, for members of team E:



Members within Team E

- Numbers of operations of all types, for members of team E:

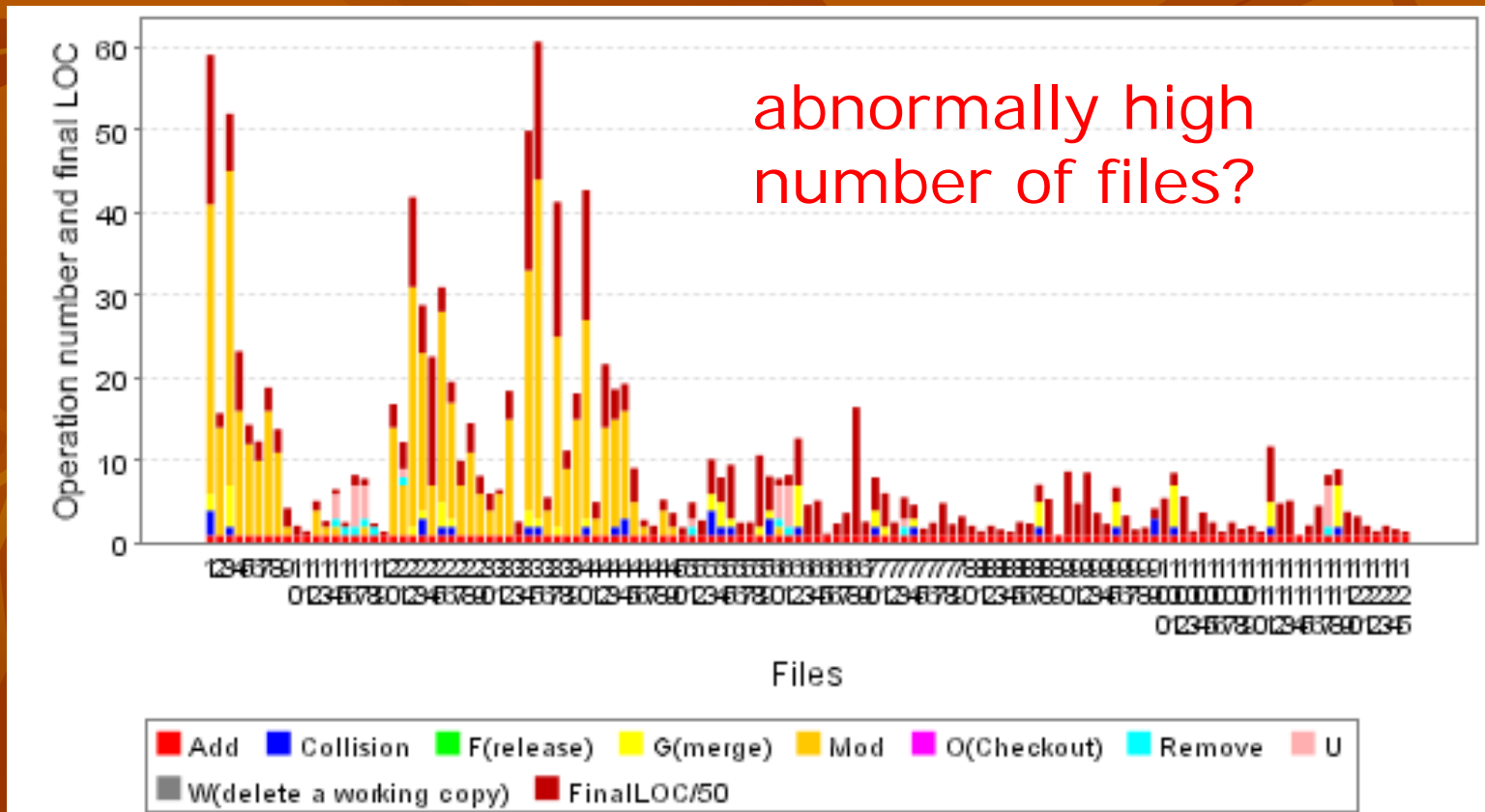


Files of a Team

- Questions:
 - How many source files does a team have?
 - Which files are changed by multiple members?
 - Which files are heavily changed?
 - Which files are central?
 - Who is the likely author of a file?
 - What is the distribution of operations?

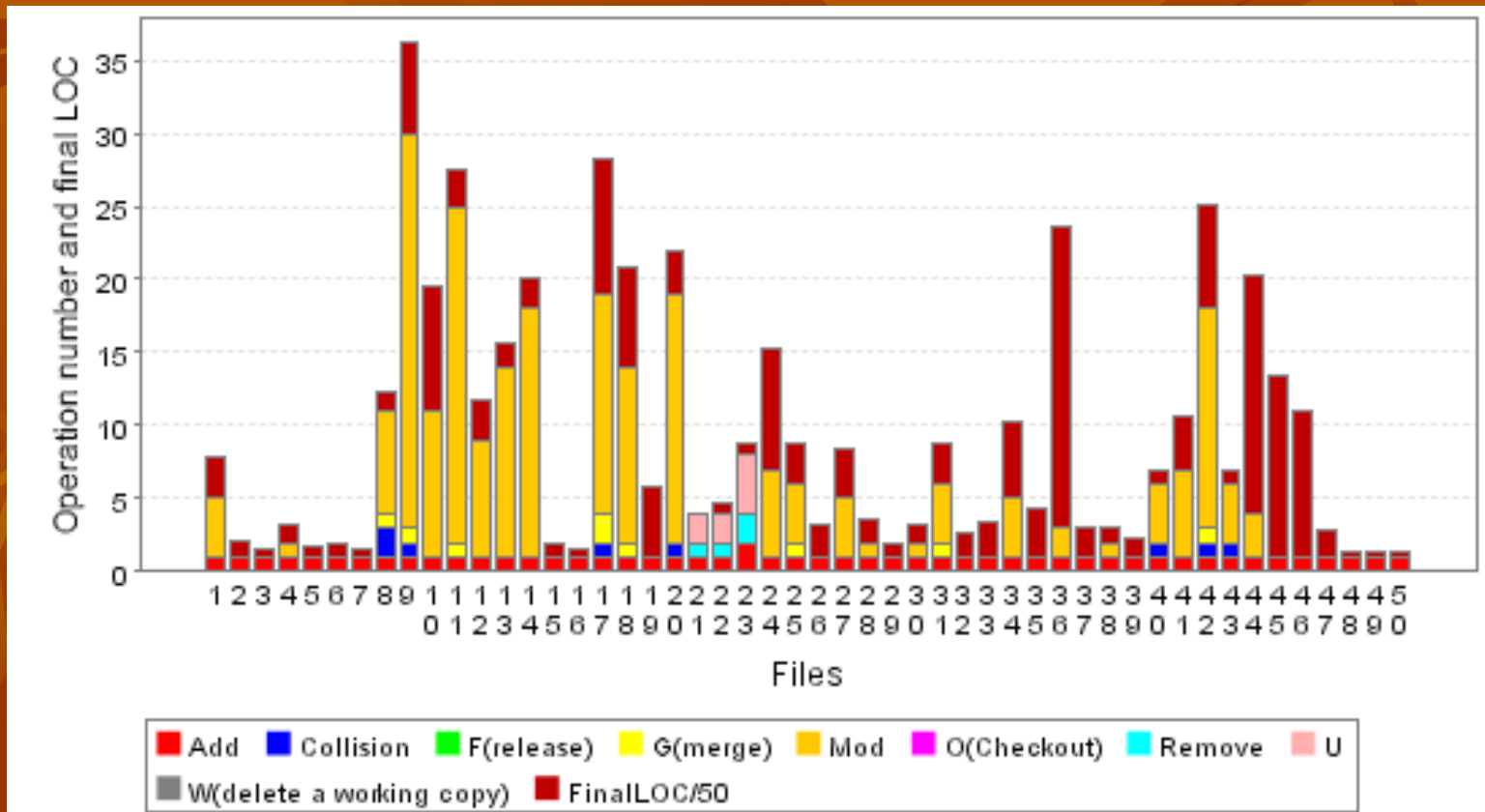
Files of Team D

- Resulting lines of code in files:



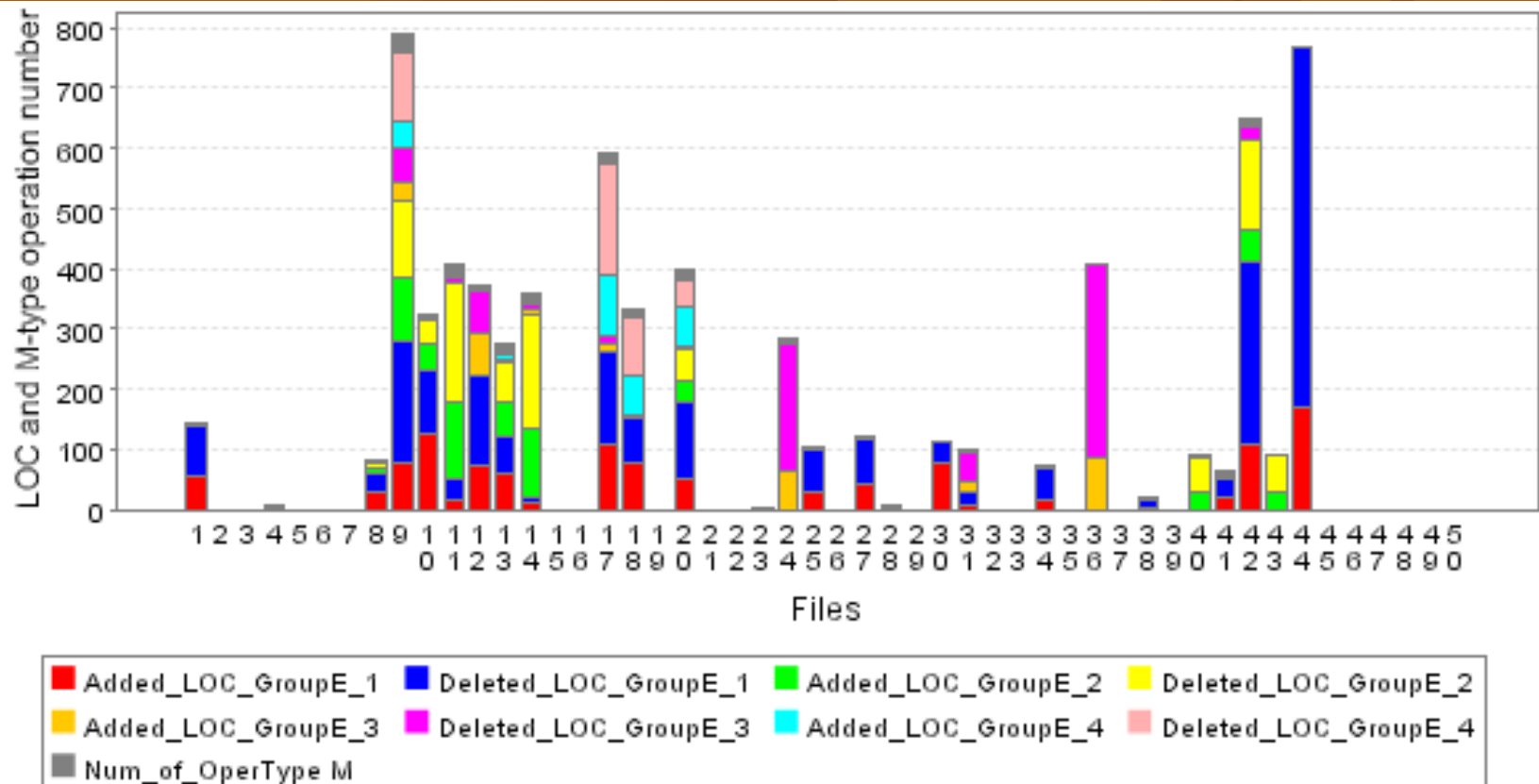
Files of Team E

- Resulting lines of code in files:



Files of Team E

- Added and deleted lines to files:



Assessments

- Issue:
 - information implicitly inferred from CVS data is revealing, but needs to be **validated**
 - it is interesting to compare it with **data explicitly provided** by developers and managers
 - these questionnaires are required for the students to complete

Team D

- Compliments:

- “because of my group members work ethics in being determined to start early, **work regularly**, and keeping each other updated on one another’s progress”
- “**communication was** open and **constant** via ICQ and email”
- “each member was more than willing, if not enthusiastic, to contribute and participate”
- “I was very impressed with other members’ willingness to help other members with problems in their ‘assigned’ areas”
- “**student 2** did a **lot of work** with the coding (especially the interface design”
- “**very impressed** with the **effort** that **student 2** and student 1 put into the GUI”

Team E

- Complaints (stage 2)
 - “some **confusion** as to **who was doing what**. Some parts were done out of order so we couldn’t do our part until all this other stuff was built”
 - “some **miscommunication** of what the plan was”
 - “concentrated largely on the front-end and the **back-end** was **poorly formed** and probably will have to be redone for the next part of the project”
- Complaints (stage 3)
 - “**reverted** to the **old ways** of the computer geek”
 - “it is better to underestimate yourself than to overestimate”

Vision

- Ongoing work:
 - complete, integrated data collection (e.g., include PSP measures, code metrics, defects, communication)
 - complementary set of “diff” algorithms (e.g., compare versions, identify refactorings and co-evolution)
 - allow teams to reflect on their own progress and process (process mentor)
 - improved visualization

Vision

- Ongoing work:
 - data-mining
 - try to find some associated operation and set reminders
 - identify developers according to roles
 - in-class experiments
 - we are looking for partners!
 - {kenw ,stroulia} @cs .ualberta.ca
 - <http://www.cs.ualberta.ca/~stroulia/jreflex/>

Discussion

- Issues:
 - granularity
 - operations versus MRs
 - interpretation
 - knowledge of underlying process
 - generalizing
 - open source? scalability?
 - conducting studies
 - ethical approval, anonymizing data, privacy
 - benchmark data sets