# Communication Problems in Global Software Development: Spotlight on a New Field of Investigation

Sébastien Cherry
Pierre N. Robillard

École Polytechnique de Montréal

# Problem / Motivations

- Some human aspects might be the source of numerous problems in software engineering
- Some researchers have observed that informal communications take up a considerable portion of the time spent on software projects
  - Perry, Staudenmayer and Votta (1994)
  - Robillard and Robillard (2000)

# Problem / Motivations (2)

- Herbsleb and Grinter (1999) support that distance in GSD raises barriers to informal communications and results in coordination problems

- No known research have tried to describe the content of these ad hoc collaborative activities

# Research Objectives

- To design a model of the ad hoc collaborative activities found in an industrial software engineering setting and characterize them as well as to describe the content of the communications that ensue

- To generate a series of hypotheses emerging from the results of this research that could later be validated by confirmatory research

# Theoretical and Practical Relevance

- Theoretical relevance:
  - To generate a series of hypotheses to create a theoretical base of knowledge in this domain
- Practical relevance:
  - To improve SE processes with new state-of-the-art rules or formal practices
  - To have some insights into the tools needed to support communications in GSD contexts

# General Approach

- Participant observation
  - Enterprise:
    - International organization in software development
    - In the sector of enterprise information systems
    - Where there is a clear defined software process
  - Software development team:
    - 8 teammates
    - 4 to 17 years of experience in software development
    - Wide range of ages and schooling
    - Variable length of service in the company

# Data Collection

- The first phase of data collection includes:
  - 185 hours of audio-video recordings of working sessions over 37 workdays
  - The capture of 2496 e-mails exchanged by the 8 teammates
  - A daily backup of the source code and other relevant artefacts

# Data Analysis

- Use of the ESDA (Exploratory Sequential Data Analysis)
- Well-suited for:
  - Exploratory researches where the objective is to find answers to research questions
  - To find patterns among the empirical data
  - Where the sequential integrity of the data must be preserved

# Data Analysis (2)

- The ESDA proposes eight data manipulations
  - "Coding" (the most important), consists of labelling each sequences of data by means of a code contained in a exhaustive, exclusive and restricted list of categories
- Iterative process of theory induction

# Preliminary Results and Observations

Legend:
- Ad hoc collaborative activities
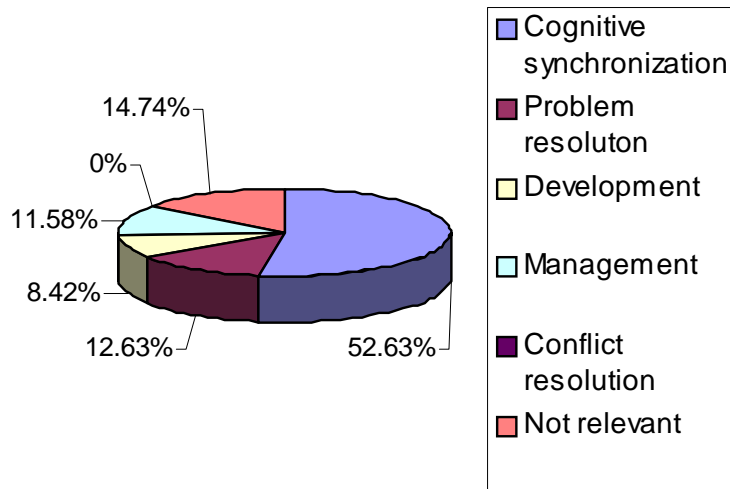- Other types of activities

49%

51%
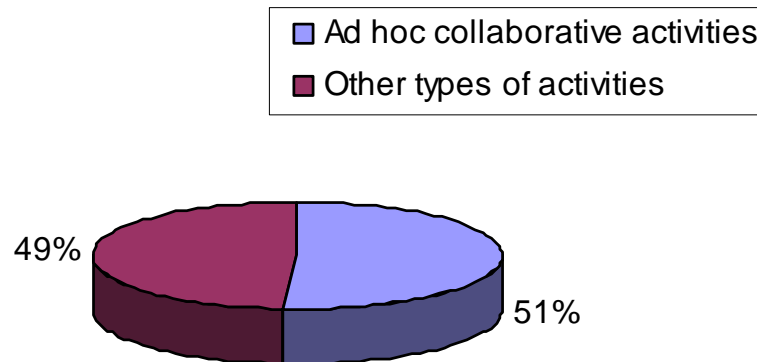
**Distribution of time spent in ad hoc collaborative activities in comparison with other types of activities**

# Preliminary Results and Observations (2)



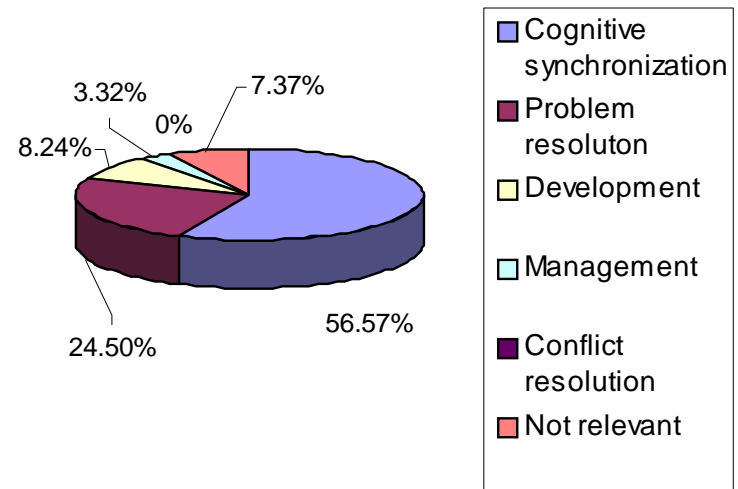Percentage of time spent on ad hoc collaborative activities
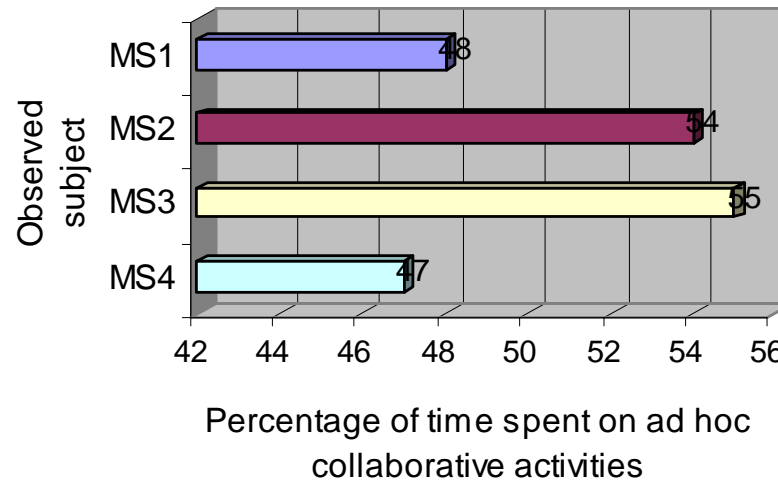
**Percentage of time spent on ad hoc collaborative activities by observed subject**

# Preliminary Results and Observations (3)



**Distribution in number of occurrences of ad hoc collaborative activities identified**

Legend:
- Cognitive synchronization
- Problem resoluton
- Development
- Management
- Conflict resolution
- Not relevant

14.74%
0%
11.58%
8.42%
12.63%
52.63%



**Distribution in terms of time spent on ad hoc collaborative activities identified**

Legend:
- Cognitive synchronization
- Problem resoluton
- Development
- Management
- Conflict resolution
- Not relevant

3.32%
7.37%
0%
8.24%
56.57%
24.50%

# **Conclusion**

- Software developers spend a lot of time in forms of informal and spontaneous collaborations
- No research found have tried to describe the content of these activities

# Conclusion (2)

- The practical objectives of this research are:
  - To propose software process enhancements with the aim of rendering the collaboration between teammates more effective
  - To obtain some insights into the tools needed to support communications in a distributed software development context
- Preliminary results tend to demonstrate that a data model and certain patterns are emerging from the vast quantity of data amassed

# Questions?