
Expertise Browser: a Quantitative Approach to Identifying Expertise

Audris Mockus

Software Technology Research Department, Avaya Labs

James Herbsleb

Bell Labs, Lucent Technologies

Audris Mockus: Expertise Browser

Outline

Background

- Research question
- Motivation: multi-site/large-scale software projects
- Approach

Estimating expertise

Expertise Browser

- Finding relevant people
- Showing personal profiles

Summary

Audris Mockus: Expertise Browser

Research Question

Can the expertise about a piece of software be

- a) empirically mapped to developers who have it?
- b) used to improve the software engineering practice?
 - 1) streamline expertise finding
 - 2) make knowledge about the project widely available

Motivation

A serious problem in large-scale software development: delays locating experts

- Typically, only a few individuals (if any) understand the entire SW system and they are in high demand
- Often the “best match” for a task (once known) is unavailable – what is the “next best match”?
- The explicit repositories of knowledge, e.g., skills database are usually too generic, too outdated, and it is unrealistic to improve that.

Approach

Present relationships between people and the code they know

- **Quantify person's experience with a part of code using change history of the code**
- **Estimate relationships between:**
 - **The structure of the code**
 - **Organizational structure**
- **Deploy a tool to allow:**
 - **Identify experts quickly and easily**
 - **Compare experts with one another**
 - **Present a profile of expertise for an expert**

Background

Software is created incrementally, via changes recorded by VCS

- **A delta is single checkin (ci/commit/edput) representing an atomic modification of a single file with following attributes**
 - **File**
 - **Date**
 - **Developer**
 - **Comment**
- **Other attributes that often can be derived:**
 - **Size (# of lines added, deleted)**
 - **Lead time (interval from start to completion)**
 - **Purpose (Fix/New)**

Expertise and Experience

Expertise: *Ability effectively to understand, enhance, fix, or test a part of a software system*

Experience: *Amount of work (number of changes) performed on a part of a software system*

Expertise increases with Experience

- **Productivity increases**
- **Quality increases**
- **Supported by interviews**

Experience Atoms (EAs)

A delta is a unit of experience or EA

- **Developer making it gained experience with that code**
- **Code unit that was changed gained “knowledge” from the person changing it**

Example experience measures

- **Coding experience of an individual or a group**
 - **Number of EAs done on a code unit/language/project**
- **Testing experience**
 - **Number of problem reports raised by a subject**

Expertise Browser

Obtains and presents relationships between code and people and organizations based on EAs shared between parts of software product and people

Deployment

Simplicity of access – most useful to participants that are hardest to reach, home workers/remote sites

Intuitive to use – to minimize training

Basic functional requirements

Handle a large project – visual representation/response time

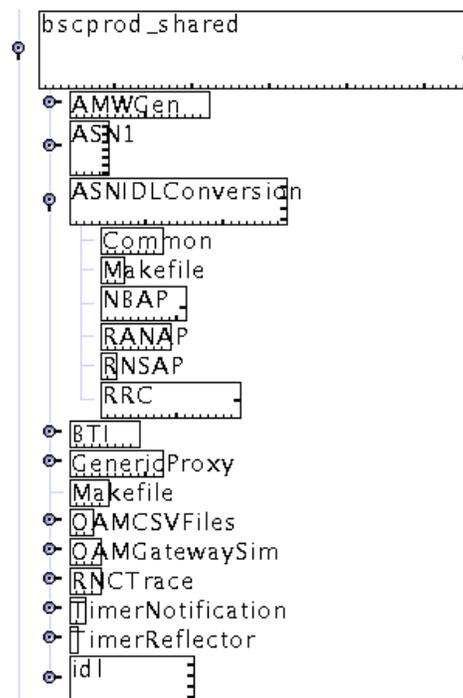
Tasks: show menu of code related expertise, org. profile

Product View

An expandable tree normalized by changes (based on directories or subsystems/modules)

Each node on the right is a module/file or a set of modules

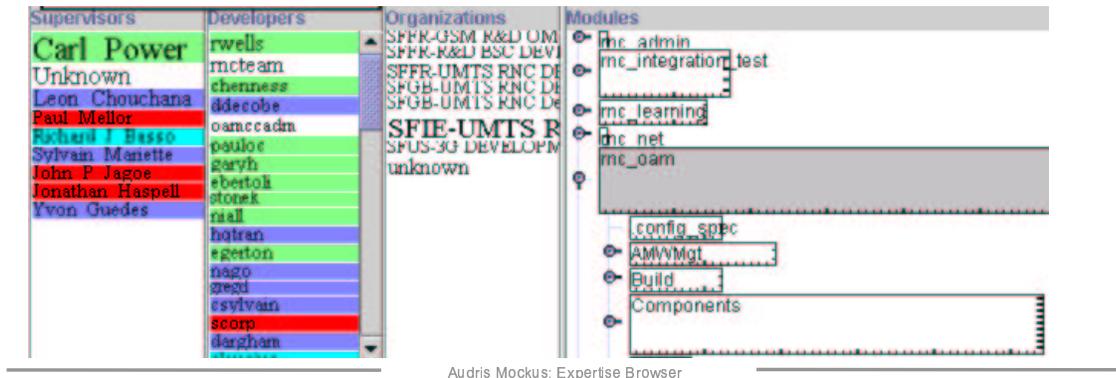
- Height - $\sqrt{\text{\#of EAs}/10} + \text{font height}$
- Width - 5 pixels per contributing subject



Task1: Expert Search

Select a code unit to show experts

- All developers, their supervisors, and organizations ordered by expertise
 - Developers at the top are most relevant
 - Largest font reflects most experience
 - Color identifies geographic location of the subject



Audris Mockus: Expertise Browser

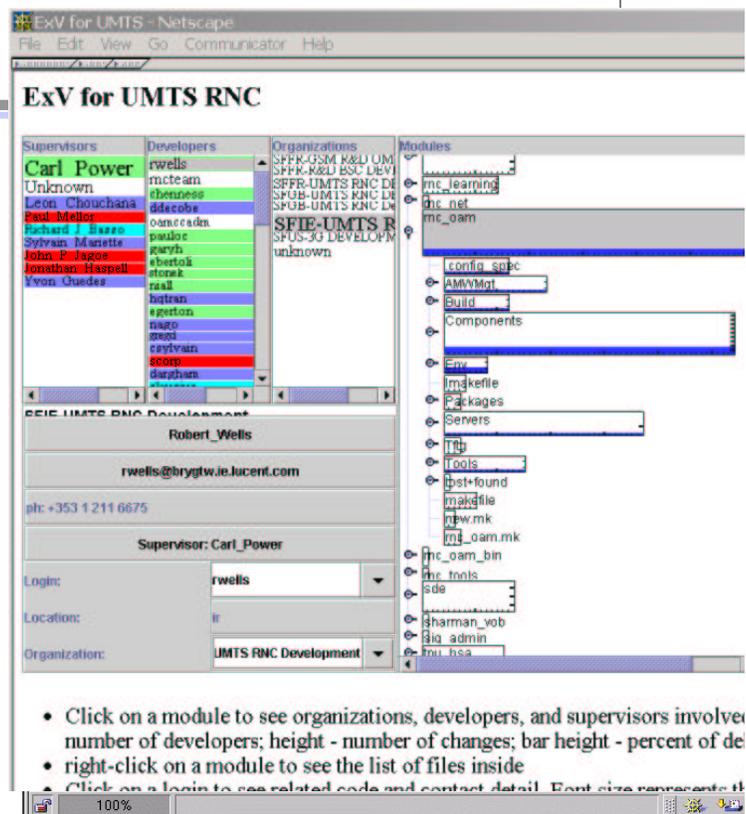
Task 2: Resume

Select a person to show

- Fraction of EAs for CUs
- Contact info

Select an org. to show

- All developers in the organization/group
- Fraction of EAs contributed by these developers for each CU



- Click on a module to see organizations, developers, and supervisors involved
- number of developers; height - number of changes; bar height - percent of dev
- right-click on a module to see the list of files inside

• Click on a login to see related code and contact detail. Font size represents th

Audris Mockus: Expertise Browser

Task 3: Work Awareness

Estimate persons “Home Area” using recent changes

Define impact measures, e.g.,

- Same line/file/module changed
- Functions called are changed

Determine/show others who do current work with potential impact

Who messed around my code?

Individual view for rwells

- Home Area: files and modules touched by rwells over last year
- Changes by others over last week on the same files and modules
- Right-click to see the diffs between your version and other versions

The screenshot displays the Expertise Browser interface. On the left, a list of actions is shown: change file, own, new file, change subs, new subs, new mod, and change mod. A list of users is displayed, with 'rwells' highlighted in green. Other users listed include rnc team, chenness, pauloc, egerton, stonek, sregd, diyang, and huchro. The main area shows a file tree for 'rwells' with folders like 'rnc_oam', '.config_spec', 'AMVMgt', 'Build', 'Components', 'Env', 'Makefile', 'Packages', 'Servers', 'Tfig', and 'Tools'. Below the file tree, there is a section for 'Supervisor: Carl_Power' with fields for 'Login: chenness', 'Location: ir', and 'Organization: S RNC Development'.

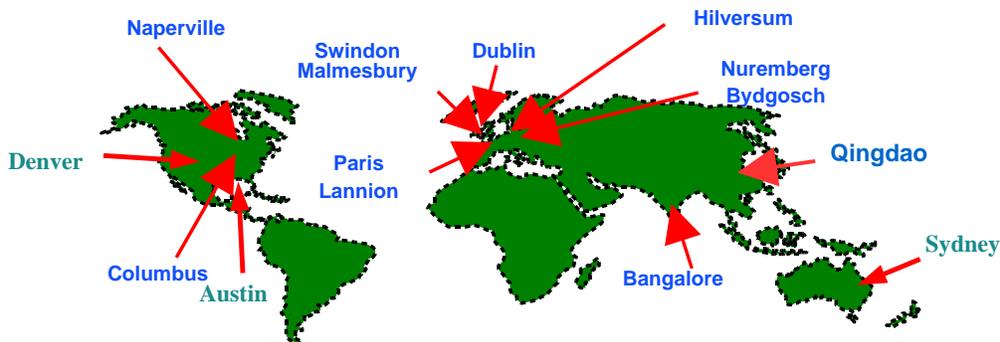
Largest Projects with ExB deployed

Wireless Call Handling (7M lines added in 200K deltas over 5 years by 110 developers, in 3 primary locations in 3 countries)

OA&M Product (6M lines added in 100K deltas over 5 years by 350 developers, 3 primary locations in 3 countries)

Wireless Base Station Controller (14M lines in 140K deltas over 3 years by 340 developers, 5 primary locations in 5 countries)

Enterprise Voice Switch (19M lines in 677K deltas over 14 years by 860 developers in 3 primary locations)



Audris Mockus: Expertise Browser

Additional observations

Use analysis

- Relatively short sessions – median 10 selections
- Different patterns of use between
 - Remote sites – focus on code related experts, more use
 - Central sites – focus on organization expertise

Work awareness

- Problem of notifying about the changes that might impact one's work
- Warning of potential physical dependencies before integration

Audris Mockus: Expertise Browser

Conclusions

The problem of finding relevant people

- Approximating experience using traces of work in ubiquitous change records

- Find experts and expertise profiles

 - compare experts

 - distribution of expertise – broad vs specialized

Challenges

- Find changes/people potentially impacting your work

 - identifying related work

 - highlighting potential conflicts

- Finding experts in requirements, architecture

- Project management applications