MSR 2005
International Workshop on Mining Software Repositories

Ahmed E. Hassan and Richard C. Holt
School of Computer Science
University of Waterloo
Waterloo, Canada

{aeehassa, holt}@plg.uwaterloo.ca

Stephan Diehl
Department of Computer Science
Catholic University Eichstätt
Germany

diehl@cs.uni-sb.de

ABSTRACT

This one-day workshop is to bring together researchers and practitioners to consider methods that use data stored in software repositories (such as source control systems, defect tracking systems, and archived project communications) to further understanding of software development practices.

1 THEME AND GOALS

Software repositories such as *source control systems*, *archived communications between project personnel*, and *defect tracking systems* are used to help manage the progress of software projects. Software practitioners and researchers are beginning to recognize the potential benefit of mining this information to support the maintenance of software systems, improve software design/reuse, and empirically validate novel ideas and techniques. Research is now proceeding to uncover the ways in which mining these repositories can help to understand software development, to support predictions about software development, and to plan various aspects of software projects.

We expect the presentations and discussions to continue on a number of general themes and challenges, from the previous workshop (MSR 2004) held at ICSE 2004, such as:

- Engineering challenges related to the infrastructure and tools needed to recover useful data from these repositories
- Methods of integrating mined data from various historical sources
- Development and validation of approaches to visualize and present such data
- Use of recovered history for system understanding and analysis of change patterns
- Modeling of defects and software reliability using data from such repositories
- Uncovering of the social processes and interaction between the development community
- Discovery of techniques to facilitate software reuse

2 SCOPE AND TOPICS OF INTEREST

The workshop's scope is the general field of software repository mining. Relevant topics include but are not limited to the following:

- Approaches to study the quality of the mined data along with guidelines to ensure the quality of the recovered data
- Proposals for exchange formats, meta-models, and infrastructure tools to facilitate the sharing of extracted data and to encourage reuse and repeatability
- Models for social and development processes that occur in large software development projects
- Search techniques to assist developers in finding suitable components for reuse
- Techniques to model reliability and defect occurrences
- Analysis of change patterns to assist in future development
- Case studies on extracting data from repositories of large long lived projects
- Suggestions for benchmarks, consisting of large software repositories, to be shared among the community

Position papers are no more than 5 pages long and are reviewed by the workshop's program committee in terms of their relevance to the aims of the workshop and their technical content. Accepted papers are posted on the workshop's web site prior to the workshop at:

```
http://msr.uwaterloo.ca
```

3 FORMAT AND EXPECTED OUTCOME

The program committee selects for presentation papers that can serve as the basis for fruitful discussions. The papers are chosen so that a broad range of stakeholders from across the software engineering discipline are represented in the workshop. An emphasis is on making the workshop interactive with many discussion slots assigned throughout the schedule. Presentation are short with strict time limits to ensure ample discussion time. In addition, there is a wrap-up session at the end of the day to set goals for further research in the area of mining software repositories.

Since MSR 2004 was the first workshop to focus on the field of mining software repositories, it was exploratory in nature with a variety of results in different areas of the field. We expect MSR 2005 to retain an exploratory feel while at the same time promoting collaboration and systematic comparison of approaches and techniques. This mixture of exploratory work with systematic comparisons and collaborations should assist in maturing the field of mining software repositories and in directing it to take a central role in supporting software development practices.