MSR Challenge 2011: Eclipse, Netbeans, Firefox, and Chrome

Adrian Schröter
University of Victoria, Canada
schadr@acm.org

ABSTRACT

The *MSR Challenge* aims at offering researchers and practitioners in the area of Mining Software Repositories a shared set of software repositories, enabling them to compare their tools and approaches. This year, the main theme of the challenge was on the comparison of projects. We selected four open source projects, and challenged participants to use their brains, tools, computational power, and magic to compare them and uncover interesting similarities and differences. The projects were Eclipse and Netbeans, two popular IDEs written in Java (Group 1) and Firefox and Chrome, two web browsers written in C/C++ (Group 2). We encouraged participants to analyze more than one project, ideally in the same group but allowed them to analyze a single project.

Categories and Subject Descriptors

D.2.m [**SOFTWARE ENGINEERING**]: Miscellaneous

General Terms

Standardization,Measurement,Experimentation

Keywords

Data Repositories, Report

1. INTRODUCTION

The Mining Challenge is a track of the Working Conference on Mining Software Repositories (MSR), which brings together researchers and practitioners who are interested in applying, comparing, and challenging their mining tools and approaches on the software repositories of open source projects.

This year, the MSR Mining Challenge focused on comparing issue tracking and code repositories of project pairs.

We carefully chose four projects[1] in a way that we gave a choice between two major programming languages, Java and C/C++. On the Java side, we offered data sets from two IDEs, Eclipse and Netbeans. Both IDEs have been analyzed through the course of MSR, and Eclipse data has been the reference data set of several past challenge tracks. Firefox and Chrome were the two C/C++ project that we offered for comparison. With Chrome only open sourced recently, there has been little research in MSR using Chrome data, but on the other hand Firefox has been offered as a data set for the MSR Challenge in 2008.

Thus, this year's Mining Challenge provided four different large data sets comprising two pairs of projects with similar purposes, Eclipse and Netbeans, and Firefox and Chrome. All reference data sets came with both a source code repository and an issue tracking repository, offering many possiblities to analyze and compare projects.

2. PREVIOUS WORK

Since 2006, the MSR Mining Challenge has brought together researchers and practitioners who are interested in applying, comparing, and challenging their mining tools and approaches on a shared set of software repositories. The MSR Mining Challenge chairs and program committee have selected a set of mining targets, and then proposed a kind of emphasis for the challenge. Each year normally consists of two ategories of challenges: prediction and general challenge.

Prediction Challenge.

The MSR literature is full of papers about predictor models based on the information obtained from software repositories. However, it is a challenge to publish and submit predictions before the events take place. Past editions proposed to forecast the number of changes and bugs for Eclipse (2007, 2008), Firefox (2008), code growth of the GNOME desktop suite (2009), and bug report growth of Debian (2010) [4]. The winner of the prediction challenge is selected based on the prediction accuracy. This year we skip the prediction challenge.

General Challenge.

The general challenge is a venue where researchers can submit any kind of short report about the selected target projects and repositories. The idea is to give useful feedback

---

[1]The challenge data can be found at http://2011.msrconf.org/msr-challenge.html

Title	Data Used
Operating System Compatibility Analysis of Eclipse and Netbeans Based on Bug Data [6]	Eclipse Bugzilla & Netbeans Bugzilla
What Topics do Firefox and Chrome Contributors Discuss? [2]	Chrome Issue Tracker & Firefox Bugzilla
A Tale of Two Browsers [1]	Chrome Issue Tracker & Firefox Bugzilla
Do Comments Explain Codes Adequately? [5]	Eclipse Code Repository & Netbeans Code Repository
Apples Vs. Oranges? An exploration of the challenges of comparing the source code of two software systems [3]	Eclipse Code Repository & Netbeans Code Repository

Table 1: Submissions and the data they analyzed.

to the authors, so they appreciate the value of MSR research for practitioners. Moreover, it allows current researchers to demonstrate state-of-the-art mining techniques on a public data set. Researchers can choose any tool and study any kind of public repository, although the MSR Challenge provides researchers with some already extracted data sets and repositories. In this manner, MSR supports future research by bootstrapping many researchers with useful data, allowing them to skip the mirroring and extraction step. The winner of the general challenge is chosen by the attendees of the workshop based on the quality of their presentations. Since 2008, winners of the MSR Challenge are awarded a prize.

3. THE CHALLENGE

The main focus of the Mining Challenge 2011 lied on comparisons. Traditionally the Mining Challenge challenged to compare skills and tools in the area of mining software repositories, but this year we took it one step further and challenged practitioners and researchers alike to not only compare their skills and tools with each other but to use these skills and tools to compare similar software projects.

The goal of the Mining Challenge 2011 was to encourage researchers and practitioners to take up the challenge of focusing on similarities and differences in similar programs. Over the course of the year, many projects were analyzed to assess different aspects or test different theories, such as what predicts failures and effort. Also many studies used a number of projects to investigate their generalizability of their methods and theories, but often stopped at that. What is missing so far are studies that systematically investigate questions investigating the difference and similarities of software projects. "Do web browsers evolve similarly?" or "Do web browsers have a similar architecture?".

4. DATA

Since this year's challenge was focusing on comparing similar projects we provided not one but two pairs of software projects. The first pair consisted of two sophisticated IDEs, whereas the second pair represented two popular web browsers.

4.1 IDEs

The first project pair consisted of the two IDEs Eclipse and Netbeans. Both IDEs are written in Java and serve the same purpose, namely supporting a developer in writing and maintaining source code.

Eclipse and Netbeans are Integrated Development Environments (IDE) developed in Java. Both enjoy great pop-

ularity among developers for the vast number of tools it provides and its extensibility through third party plug-ins. Some of the standard tools Eclipse and Netbeans provide a developer with are back ground compiling, a debugging tool set and connectors to popular version control systems.

We chose Eclipse and Netbeans for several reasons: (1) IDEs have been the subject of many studies in the past, which could serve as inspiration for a comparison study. (2) IDEs are programs that have great relevance to practitioners and have often been used by our community to bring research knowledge in the form of plug-ins to them. (3) Eclipse and Netbeans have long reached a maturity which provides us with a rich development history.

We collected their code repository and issue tracker data. The Eclipse foundation provided a downloadable dump of their CVS code repository. For Netbeans we mirrored their Mercurial version control systems repository using the clone functionality provided by Mercurial. To gain access to the issue tracker we asked the Eclipse foundation and Netbeans foundation if they could provide us with a dump of their Bugzilla database. They provided us with a dump of the Bugzilla database that excludes personal information such as email addresses or user's real names.

4.2 Browsers

The other pair of projects consists of the two web browsers Firefox and Chrome. Firefox and Chrome represent two of the most popular web browsers in use. Although Firefox has been around for 3 years longer than Chrome, Chrome gained on popularity ever since its first release. Both browsers provide similar features ranging from displaying HTML pages and running Javascript and extension points to develop plug-ins.

As with the two IDEs we chose Firefox and Chrome because they have both reached a mature state with a sufficiently large development history. Besides that, both projects enjoy great popularity, they share a number of core functionalities, and features that greatly enhances comparability of the two projects. Furthermore, both browsers are developed using C/C++ which also makes it easier to compare them as well as giving a good alternative to the two Java IDEs of the other project pair.

For both browsers we obtained a copy of their version archives. Firefox used to use CVS as their main repository which we mirrored using rsync[2] and started to use Mercurial. Chrome uses a centralized SVN repository that can easily be mirrored. For issue tracking Firefox uses Bugzilla and Chrome uses a Google issue tracker. We could obtain

---

[2]https://wiki.mozilla.org/How_to_Create_a_CVS_Mirror

neither a dump of the Firefox or Chrome issue tracker due to security bugs that they do not want to be publicly available, but we could obtain all publicly available issue reports.

5. SUBMISSIONS

We received a total of six submission of which we accepted five. The submissions presented a good mix between analyses of the browsers and the IDEs. Table 1 gives an overview of the submissions we accepted and the data sources they used. Following we summarize the findings of our accepted submissions.

The two studies investigating the issue tracker repositories of Eclipse and Netbeans found that Eclipse outperformed Netbeans with respect to comment quality and stability on Windows, Linux and MacOS. Comments in Eclipse were more accurate in describing the actual issue than in Netbeans [5]. Moreover, Eclipse had a stable number of operating system related issues over time, whereas Netbeans exhibited an increasing trend in the number of operating system related issues [6]. Though the numbers suggested that Eclipse has less operating system related issues there could be alternative reasons as to why the numbers differed other than the quality of the code, such as Eclipse users being less picky than Netbeans users.

Firefox and Chrome both popular exhibited two distinct profiles. Firefox presented the more mature project with rigorous processes and a long release cycle. On the other hand Chrome being the newer project evolved more rapidly by a very enthusiastic user base that pushed for new releases at the price of stability [1]. The topics discussed in bug reports also reflected this difference between the two user-bases' interests. In Firefox some of the most popular topics related to low-level issues, such as the product building process, and general layout issues whereas Chrome issues contained discussions about optimizations and multimedia use. The topics discussed in both browser communities were not completely distinct for example both communities discussed HTML related topics such as page layout [2].

German et al. [3] pointed out that we need to be aware of certain pitfalls in all comparison studies, for example code repositories often contain large quantities of code that is not needed to build the product, such as test suits. These pitfalls can introduce a major bias into research results of comparative studies and might even invalidate findings.

6. CONCLUSIONS

The MSR Challenge continues to supply the community around mining software repositories as well as other empirical researchers in the course of their research with data. This year we provided data of four major open source software projects: Eclipse, Netbeans, Firefox, and Chrome. We attracted a number of high-quality submissions that together covered both project pairs and generating interesting insights about the differences and similarities of Eclipse and Netbeans as well as Firefox and Chrome.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] O. Baysal, I. Davis, and M. W. Godfrey. A Tale of Two Browsers. In *8th Working Conference on Mining Software Repositories (MSR)*, New York, NY, USA, May 2011. ACM.

[2] M. L. Bernardi, C. Sementa, Q. Zagarese, D. Distante, and M. D. Penta. What Topics do Firefox and Chrome Contributors Discuss? In *8th Working Conference on Mining Software Repositories (MSR)*, New York, NY, USA, May 2011. ACM.

[3] D. M. German and J. Davies. Apples Vs. Oranges? An exploration of the challenges of comparing the source code of two software systems. In *8th Working Conference on Mining Software Repositories (MSR)*, New York, NY, USA, May 2011. ACM.

[4] A. Hindle, I. Herraiz, E. Shihab, and Z. M. Jiang. Mining challenge 2010: Freebsd, gnome desktop and debian/ubuntu. In *7th Working Conference on Mining Software Repositories (MSR)*, pages 82 –85. IEEE, May 2010.

[5] Y. Hirata and O. Mizuno. Do Comments Explain Codes Adequately? In *8th Working Conference on Mining Software Repositories (MSR)*, New York, NY, USA, May 2011. ACM.

[6] X. Wang, E. Baik, and P. T. Devanbu. Operating System Compatibility Analysis of Eclipse and Netbeans Based on Bug Data. In *8th Working Conference on Mining Software Repositories (MSR)*, New York, NY, USA, May 2011. ACM.