

Chapter 4

Open Source Projects Recommendation on GitHub

Mohamed Guendouz

Dr. Moulay Tahar University of Saïda, Algeria

Abdelmalek Amine

Dr. Moulay Tahar University of Saïda, Algeria

Reda Mohamed Hamou

Dr. Moulay Tahar University of Saïda, Algeria

ABSTRACT

This chapter discusses the design and the implementation of a recommender system for open source projects on GitHub using the collaborative-filtering approach. Having such a system can be helpful for many developers, especially those who search for a particular project based on their interests. It can also reduce searching time and make search results more relevant. The system presented in this chapter was evaluated on a real-world dataset and using various evaluation metrics. Results obtained from these experiments are very promising. The authors found that their recommender system can reach better precision and recall accuracy.

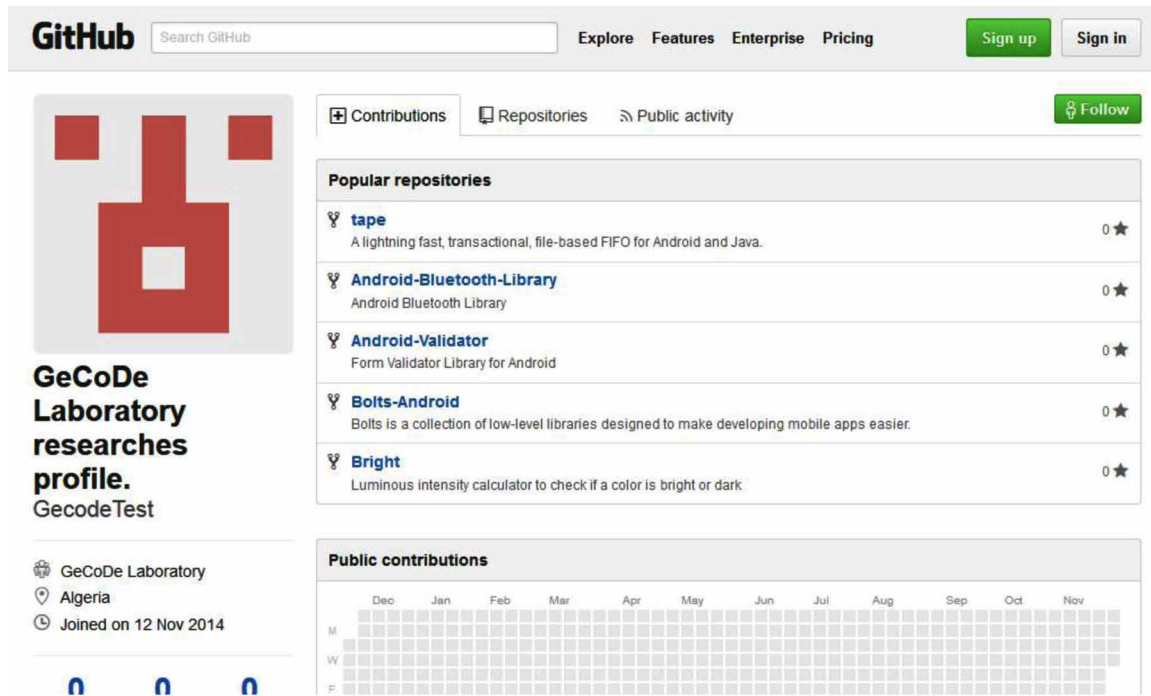
INTRODUCTION

GitHub is a very popular crowdfunding software development platform, a social coding platform and a web based Git repository hosting service, allowing anyone to participate in open source project documentation, design, coding and testing in a social way. In order to participate in these activities, a developer must create an account, allowing him to share his own projects, forking other's projects and following other developers, Figure 1 shows a sample GitHub profile.

One of the most helpful implemented features on GitHub is the fork feature, which means making a full copy of the repository of the original project. Forking a repository allows the developer to freely experiment with the project without affecting the original copy, forking is considered as the first task

DOI: 10.4018/978-1-5225-5314-4.ch004

Figure 1. Example of a GitHub profile page



to do in order to make contributions to an existing project. Another implemented feature is the Star feature, when a developer gives a star to a repository it means that he is interested in this project. For example, a developer who is interested in mobile game development may give stars to some 2d mobile game libraries like: AndEngine, LibGDX, cocos2d-x and others.

Developers are always searching for good open source projects to make project prototypes or to enhance their own software projects with new features, GitHub provides them a search functionality to do this manually without any automatic recommendations provided, Figure 2 shows a sample search page. However, searching for suitable repositories can be a difficult task and may take a long time, it can also interpret the development process of a project, for that reason the existence of an automatic recommender system for GitHub repositories may be very helpful for developers to reduce search time and make search results more relevant and organized, these are the main benefits of such a system for all developers. However, developers may benefit differently from it according to their profile type and their professional skills, for instance: a professional developer is probably searching for new programming challenges or even for business opportunities, while a beginner is probably looking for good stuff to learn something new or to improve its skills, or he is simply searching for repositories to work on. The issue that arises in these cases is how we can find a relevant content on GitHub and recommend it to a user.

In this paper, the authors present a new system for recommending relevant GitHub repositories for developers; they use a collaborative-filtering approach and they model the user behaviors as a User-Item matrix so they can apply different recommendation methods like calculating similarities between users (developers) and items (repositories) and so on. Then, the authors evaluate their recommender system on a real data set using well-known evaluation metrics, the design and the implementation of this system will be discussed in detail in later sections.

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/open-source-projects-recommendation-on-github/197107

Related Content

Minimization of Energy in Smart Phone Application Development Using Code Analysis

K. G. Srinivasa, Srinidhi Hiriyannaiah and G. M. Siddesh (2018). *International Journal of Open Source Software and Processes* (pp. 48-60).

www.irma-international.org/article/minimization-of-energy-in-smart-phone-application-development-using-code-analysis/217414

Creating Open Source Lecture Materials: A Guide to Trends, Technologies, and Approaches in the Information Sciences

William H. Hsu (2015). *Open Source Technology: Concepts, Methodologies, Tools, and Applications* (pp. 336-363).

www.irma-international.org/chapter/creating-open-source-lecture-materials/120924

Open E-Resources in Libraries

Vesna Injac-Malbaša (2015). *Open Source Technology: Concepts, Methodologies, Tools, and Applications* (pp. 133-160).

www.irma-international.org/chapter/open-e-resources-in-libraries/120911

Efficient Algorithms for Cleaning and Indexing of Graph data

Santhosh Kumar D. K. and Demain Antony DMello (2020). *International Journal of Open Source Software and Processes* (pp. 1-19).

www.irma-international.org/article/efficient-algorithms-for-cleaning-and-indexing-of-graph-data/264482

Social Networks and Internet Communities in the Field of Geographic Information and Their Role in Open Data Government Initiatives

Paula Díaz and Joan Masó (2015). *Open Source Technology: Concepts, Methodologies, Tools, and Applications* (pp. 1586-1618).

www.irma-international.org/chapter/social-networks-and-internet-communities-in-the-field-of-geographic-information-and-their-role-in-open-data-government-initiatives/120990