Open Source Communities in China (Mainland): An Overview

Yi Wang¹, Fan Li², and Jiguang Song³

¹ Dept.of Computer Science and Engineering, Shanghai Jiaotong University, 200240 Shanghai, P.R.China yi_wang@sjtu.edu.cn
² School of Management, Zhejiang University (Zijingang), 310058 Hangzhou, P.R.China shelleylf@gmail.com
³ School of Software Engineering, Beihang University, 100083 Beijing, P.R.China aiquanshui@126.com

Abstract. Open Source Software development has been an important software development way in last decade. Meanwhile, the Open Source Communities also grow fast in these years and become more and more important. This paper presents a preliminary study to China's Open Source Communities. We provide some descriptive results according our survey and point out some problems threatening the further developments of China's Open Source Communities. We also provide a model for assessing the maturity of the Open source community. A case study about one typical China Open Source Community also provided.

1 Introduction

1.1 Motivation

In recent several years, open source movement and community-driven software development have been popular in software industry. Many of these projects won great success, for example Apache [1], Ubuntu [2] [3] and so on. In China, open source development also has been a hot topic. It has also been attracted many scholars' interest [4]. However, quite a few studies has focused on these online communities that driven these software developments. This paper provides some preliminary research in this area.

Studying the properties of Open Source Communities is important. It can provide better understanding of community-driven software development to the open source software developers. It is helpful in building a vibrant community, and facilitates efficient communications within the community. And ultimately, increases the productivity of software and provides more reliable and usable software. Besides, for Open Source Communities in China are mostly in their startup phase, researches on this area can help them a lot in their further developments.

We first give a brief introduction of our data source and Categories of Open Source Communities in China in Section 1.2 and 1.3. Then a descriptive analysis of the communities in our survey is presented in Section 2. A preliminary model to assess the Open Source Communities' maturity is given in Section 3. Maturity Index (MI) also defined in this section. Section 4 is a case study of a typical China Open Source Community (Woodpecker Pythonic). Related works and Conclusion are described in Section 5 and 6 respectively. The future work is included in section 7.

1.2 Data Source

The data of this research mainly comes from three sources:

- Questionnaires mailed to 300 members of 10 Open Source Communities (listed in the end of this paper) in China. And then, we got 216 replies.
- Open materials available on the Internet.
- Related data in some projects' bug report archives, email archives, and the CVS repositories.
- The authors' personal experiences as Open Source Community leaders.

1.3 Categories of Open Source Communities in China

The Open Source Communities in China can be divided into three categories, they are: (1) Branches of international communities, 2 in our survey, for example, Ubuntu China. (2) Local Open Source Communities, 8 in our survey, for example, Woodpecker. In this paper, the focus is mainly on the latter category. The reason is that branches of international communities are more influence by their superior communities that do not locate in China.

2 Descriptive Analysis

We first provide descriptive analysis of the Open Source Communities in China according the data we gathered. The descriptive analysis can bring the direct understanding of current China's Open Source Communities. We discuss the Venues, structure, governance, financial support, and license selection of China's Open Source Communities in this section.

2.1 Venues

Venues refer to the places where communities' activities take place. Any online communities, no matter their sizes, must have at least one regular venue, open source communities also are not exceptions. However, Open Source Communities are not *Pure* online communities, some activities also occurs in our real life not only just in the virtual digital world. We discuss these two kinds' venues respectively in the following.

2.1.1 Online Venues

For the Open Source Community's members do not live in one location, they can not meet other members as soon as they want. The online venues satisfy this demand of instant communications. Typical online Venues include mailing lists, web forums, instant message software (IM), IRC channels, and so on. Figure 1 describes the use of these venues in China's Open Source Communities.



Fig. 1. The use of different online venues

From figure 1, we can find that China's Open Source Communities' members are more like to use web forums and instant message software while mailing lists and IRC are extremely popular in Open Source Communities have European and North American backgrounds.

2.1.2 Offline Venues

Offline Venues vary a lot according their scale, from several persons meetings to nation-wide conferences. For example, Ubuntu (A typical worldwide Open Source Community) holds several worldwide conferences and organize lots of small workshops. However, Open Communities in China seldom hold large-scale conferences. In our survey, no one of these 8 local communities has held this kind conference. Besides, most of them even do not plan to do so. The chief reason is the lack of necessary expenditure. However, there are some regular offline meeting among developers living in the same city and these core developers.

2.2 Structure

Every member in the community has his/her role, so we can build the community's structure according the members' role. The structure of Open Source Community is very important because of the structure shows out the composition of the community's members. Compared to other online communities in other forms, the Open Source Communities need more contributors. A health Open Source Community should be in Onion-Shaped Structure, while an unhealthy one is not.

2.2.1 Onion-Shaped Structure

Here we use the roles defined by Crowston and Howison [5]. The members of Open source community can be divided as follow:

1. *Core developers* (members): This group can be quite small—three to ten is adequate for most projects. This group includes founders, project leaders and release coordinators. Every person in this group is familiar with others. Their main task is decision-making.

2. *Co-developers* (members): They are people who provide code for review by core developers; they are the main force in development process.

3. *Active users*: These persons contribute by testing new releases, posting bug reports, writing documentation, and, provide help to passive users—those who use the code without contributing themselves.

4. *Passive users*: They are persons who use software without any contributions to the project.



Fig. 2. The Onion-Shaped Structure is consisted by *core developers, codevelopers, and active users*. To keep a community's health, the numbers of active users is essential.

We show the Onion-Shaped Structure in figure 2. In this figure, we point out the role of different members and describe their relations. From this figure, we can find that the circle represents *active users* should be wide enough.

2.2.2 Structure of Communities in Current China

In our survey, most of intervi-ewees admit their communities in lack of active users. So they can not get enough testing and bug reports for their software. In some degree, the structure looks like an Onion loses its Scarfskin.

The reason is not the lack of users but most users download the software are silent users. For example, we have tracked an email client-tool (it has an obvious bug) for six months in an experiment. It was downloaded over 300 times. However, we get only 7 reports of this bug till now. This experiment proves that communities in China need to pay more efforts in building their own active users groups.

2.3 Governance

The Open Source Community looks like a human society in some sense. The governance of it also follows the two ways of our real world. One is the Centralized while another is the Democratic. The table 1 describes these two different governance models.

We are surprising by that the centralized governance is dominant in most communities in our survey. In the 8 local communities, only one chooses the democratic governance. We think this may result from the tradition of Chinese politics. The explicit reasons need further investigation.

	The Democratic	The Centralized
Leadership	By election	Not appointed by the public
Important decisions	Make by the whole community	Make by a small group or some individuals
Advantages	Better engagement, The decisions may be more rational, Better atmosphere in the Community. Better intendance to leaders' personal actions.	High efficiency
Disadvantages	Some times low efficiency	Limits the engagement, The decisions may cause great disagreement, No or little intendance to leaders' personal actions

Table 1. The comparison between the democratic governance and the centrilized governance

2.4 Financial Supports

Considering Open Source movement as communistic dream is nothing more than superficial [6]. Financial assistant is essential for any kinds of software development. Some communities affiliated with enterprises while others are pure voluntary organizations. The former do not need to care about money. But things are totally different for the latter. They must raise enough money for their survival and development.

Domestic enterprises have little interests in providing financial supports for Open Source Communities. In our survey, 8 local communities are all in lack of money in some degree. Only one (WiseReal) has a company as its perpetual supporter. Two of them even have to rely on members' personal donations and their financial states are extremely bad. As we mentioned in section 3.1.2, many communities can not hold offline activities for lacking of money.

2.5 Licenses

Source license models fall into three general categories [7]:

- 1. Free: the program can be freely modified and redistributed;
- 2. Copyleft: the owner gives up intellectual property and private licensing;
- 3. GPL-compatible: licenses are legally linked to the GPL licensing structure.

In practice, there are exists hundreds models of licenses. The 10 Open Source Communities in our survey mainly use 5 License models. We summarize them in the table 2, the last column show the number of communities use the license respectively.

License	Free software	Open source	Copyleft	GPL-com	Num. of use
GPL	Yes	Yes	Yes	Yes	2
LGPL	Yes	Yes	Partial	Yes	3
CCPL	Yes	Yes	Partial	Yes	1
M/NPL	Yes	Yes	No	No	1
BSD	Yes	Yes	No	No	3

Table 2. The Use and basic information about the Open Source License

2.6 Summary

Based on above descriptive analysis, we can easily find out the problems faced by most China's Open Source Communities. These are:

- 1. Lack offline activities as additions for online activities;
- 2. Do not provide support of other languages;
- 3. The structure is not reasonable for the lack of active users;
- 4. The governance mechanism needs some reformations;
- 5. Lack necessary financial supports, enterprises are not interested in supporting Open Source Communities in finance.

3 China's Open Source Communities' Life Cycle

Communities experience several phase before their totally maturity. These are: *Startup Phase, Grow Phase and Mature Phase* [8]. Therefore, to a great extent, describing the life cycle of the community equals to defining its maturity degree.

3.1 Maturity Index (MI)

For a specified Open Source Community, the paper presents a model to assess its maturity. We define Maturity Index (MI) to describe a community maturity degree. Each MI corresponds with a series of conditions as table 3. For a community, if its MI is 3, it must at least fulfill the conditions MI3 corresponds.

Table 3. The MI assessment system. MI has four grades, each related to some conditions we specifies in the right of table 3.

MI	Conditions Must Be Fulfilled
1	Has been found with some goals;
	Only do a few simple things;
	Has some core members.
2	Has built its own online infrastructure;
	Has formed its fundamental guidelines;
	Has attracted some members, but these members do not have explicit
	roles;
	Has basic governances;
	Has regular meetings within core members;
	Do not have adequate funds.

Table	3.	(Continued)
1 ante	•••	(Continuation)

2	Easy to access for potential contributors and users:
3	Easy to access for potential contributors and users,
	Has systemized community and project guidelines;
	Has enough members with different roles;
	Has regular meetings within core members and other developers;
	The communications between members are frequent;
	In lack of financial supports, but not severe;
	The community's governance is in working order;
4	Easy to access for potential contributors and users;
	Easy to use by members;
	Has a systemized and comprehensive community and project
	guidelines:
	The members have formed an Onion-Shaped community according
	their different roles:
	Has many efficient means for members to communicate in public and
	nrivate
	Has regular meeting in different levers. For example, within or
	has regular meeting in different revers. For example, within or
	between teams and the whole community;
	The community runs smooth with efficient governance;
	Has stable financial supports, can raise enough money;
	Has some activities offline;
	Has detailed plans for future development.

3.2 MI of the Communities in Our Survey

According this model, no one of these 8 Local communities' MI fulfills the standard of MI4. Following table shows the distribution of their MIs:

MI	Num. of Comm.
1	One
2	Two
3	Five
4	Zero

Table 4. MI of the Communities in Our Survey

3.3 Summary

We are regretted to find that no one can fulfill the standard of MI4. This means there are quite a few Open Source Community can be considered as a *mature* community in China. These communities still need to improve themselves in many aspects.

4 Case Studies

We examine a typical China Open Source Community in this section. One author of this paper is among the founders of it. So we can get precise and abundance information about it. This is Woodpecker Pythonic [9], which is an Open Source Community trying to build applications using Python (A popular programming Language in open source world).

Following table show some basic information about Woodpecker Pythonic:

 Table 5. Some basic information about the Woodpecker Pythonic community, includes its goal, main venues, and so on

Founded Date	June 2004
Main Goal	Python based software development
Main Venues	Web forum, IM, Wikis
Languages	Chinese (simplified)
No. of members	Over 1000 registered users, but the real contributors less than 100
Role of members	Explicit, but in lack of active users
Funded by	Members' donation, Training and service fee
Financial Status	Not very good
Governance	Centralized.
Licenses	LGPL
Document	Code of Conduct and some others
Meeting	Regular meeting on the Internet

According the Maturity Assess Model we provide in above section, the MI of Woodpecker Pythonic community is 3. The reason are: (1) they have venues for users to engage and make their contribution easily, (2) they have some formal document, (3) they have lot of users and their roles is explicit, (4) they have some channels to get financial support. Besides, the lack of active users and abundance money make it can not be a MI4 community.

5 Related Works

Till now the research on this area is extremely limited. Most Open Source researchers' interests are topics related to the software development process but not the topics related to the communities. Crowston and Howison [5] did some exploration on this filed. They defined the role of Open Source Community's members and provide a health model for assess a community's health. Our work is mainly focus on some basic statistic information about the China's Open Source Communities and builds a model to assess their maturity in order to describe their life cycle.

6 Concluding Remarks

The paper provides some preliminary research results of China's Open Source Communities. We can find that these Communities are not very mature and face many problems. This paper points out some of this problems through the survey we conducted. This paper also provides a model to assess an Open Source Community's maturity. The model use MI to describe the life cycle of a community. Each MI corresponds to a series conditions. At last, the paper also provides a brief case study.

7 Future Work

In future, we want to refine our works in following aspects:

- 1. Do more empirical studies about China's Open Source Communities;
- 2. Refine our assess model to make it more reasonable and comprehensive;
- 3. Build a dynamic model for Open Source Community's development process for guiding China's Open Source Communities' development;
- 4. Extend our models to the Open Source Community not just mainly located in China.
- 5. Provide efficient solutions for the problems faced by the Open Source Communities.

Acknowledgements. The authors thank the help of these 10 communities Members. They also want to show their appreciates to following persons: they are Miss Peng Qiu, Mr Bing Bai, Mr Jun Xie, Mr Defeng Guo and Mr Min Zhang for their constructive suggestions. We are also grateful to the anonymous referees of the extend abstract of this paper.

References

- 1. Apache Home Page http://www.apache.org/
- 2. Ubuntu Home Page http://www.ubuntu.org/
- 3. Mako Hill, B., et al.: The Official Ubuntu Book. Prentice Hall, Englewood Cliffs (2006)
- Bezroukov, N.: Open Source Software Development as a Special Kind of Academic Research http://firstmonday.org/issues/issue4_10/bezroukov
- Crowston, Howison.: Assessing the Health of Open Source Communities, IEEE computer, pp. 89–91 (May 2006)
- 6. Glass, R.L.: A Look at the Economics of Open Source: Is open source the future of the software field or a passing fad? Communications of The. ACM 47(2), 25–27 (2004)
- Wei Wu, M., Dar Lin, Y.: Open Source Software Development: An Overview, IEEE computer, pp. 33–38 (June 2001)
- 8. West, J., O' Mahony, S.: Contrasting Community Building in Sponsored and Community Founded Open Source Projects. In: Proceedings of the 38th Hawaii International Conference on System Sciences (2005)
- 9. Woodpecker Pythonic Home Page http://www.woodpecker.org.cn

Appendix: The 10 Open Source Communities in Our Survey

1 Branches of International Communities

Ubuntu Chinese Team: http://forum.ubuntu.org.cn/ Debian Chinese: http://www.debian.org/index.zh-cn.html

2 Local Open Source Communities

C3CRM: http://www.c3crm.com/ Extmail: http://www.extmail.org/ JiuJie: http://www.inlsd.org/ Linuxfans Club: http://www.linuxfans.org/ LUPA: http://www.lupaworld.com/ Rnby Chinese: http://rubycn.ce-lab.net/ Udclub China: http://cn.udclub.com/ WiseReal: http://www.wisereal.com.cn/bmb/ Woodpecker Pythonic: http://www.woodpecker.org.cn/