

複数のサブコミュニティを有する OSS コミュニティにおけるコーディネータの分析

伊原 彰 紀^{†1} 前 島 弘 敬^{†1} 梶 本 真 佑^{†1}
亀 井 靖 高^{†1} 大 平 雅 雄^{†1} 松 本 健 一^{†1}

Open Source Software (OSS) 開発は開発者やユーザなどの参加者同士の協調作業によって成り立つことから、開発者とユーザ間の情報伝達や調整行動を務めるコーディネータが重要な役割を果たしている。先行研究では、ユーザと開発者を橋渡しするコーディネータについて分析を行ったが、信頼性の高いソフトウェアを実現するために重要となるバグ報告者とコーディネータとの関係については、明らかになっていない。本稿では、開発者、ユーザ、バグ報告者の3つのサブコミュニティを介するコーディネータについて分析を行う。分析の結果、PostgreSQL では、3つのコミュニティを媒介するコーディネータの人数が増加傾向にあることがわかった。

Analysis of Coordinators in an OSS Community with Multiple Sub Communities

AKINORI IHARA,^{†1} HIROTAKA MAESHIMA,^{†1}
SHINSUKE MATSUMOTO,^{†1} YASUTAKA KAMEI,^{†1} OHIRA MASAO^{†1}
and MATSUMOTO KEN-ICHI^{†1}

Since Open Source Software (OSS) development heavily depends on collaborations among participants such as developers and users, one of the most important aspects for OSS development is the existence of coordinators who play the roles of sharing information and arranging communications among developers and users. Although previous studies have explored coordinators in OSS communities, it is not clear about the relationship between coordinators and bug reporters who are also important to realize software with high reliability. In this paper, we analyzed coordinators who are intermediate among three sub communities (developers, users, and bug reporters). As the results of analysis, we have found that the coordinators who are intermediate among three sub communities were on the increase in PostgreSQL.

1. はじめに

Linux や Apache に代表されるように、近年、Open Source Software (OSS) は商用ソフトウェアに劣らない品質と機能を備えるようになっており、社会基盤を支えるものも少なくない。たとえば、2007年6月現在、Web サーバソフトウェアの市場シェアは、Apache が世界の約7割と大多数を占めている⁸⁾。このような OSS を開発したコミュニティ (OSS コミュニティ) を分析対象として、その成功要因を明らかにするための研究が現在盛んに行われている^{4),6),7),11)}。

我々の先行研究¹²⁾では、3つの OSS コミュニティを対象として、開発者コミュニティとユーザコミュニ

ティ間で形成されるコミュニケーション構造について分析を行った。分析の結果、OSS コミュニティが活発に活動するためには開発者コミュニティとユーザコミュニティ間での橋渡し役を務め、開発者とユーザの協調作業を支えるコーディネータの存在が重要であることが明らかとなった。

しかし、一般に OSS コミュニティの参加者は開発者とユーザに限らない。特に、欠陥が少なく信頼性の高いソフトウェアの実現のためには、バグを発見し開発者へ伝えるバグ報告者が重要な役割を占める^{9),11)}。

そこで、本稿では PostgreSQL を対象として、3つのサブコミュニティ (開発者コミュニティ、ユーザコミュニティ、バグ報告者コミュニティ) に存在するコーディネータに着目し、コミュニティ成立初期からのコミュニティの成長の様子を観察したケーススタディについて報告を行う。PostgreSQL は関係データベース

^{†1} 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology

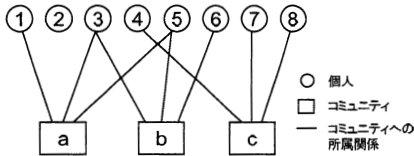


図1 アフィリエーション・ネットワーク・モデル

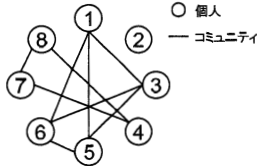


図2 個人間ネットワークモデル

管理システムであり、高い信頼性を確保する必要があるという特徴から、特にバグ報告者が重要であると考えられる。

以降、2章で関連研究について説明し、3章で本稿で想定するコミュニケーションネットワークとコーディネータについて述べ、4章で分析方法について述べる。5章で PostgreSQL を対象としたケーススタディについて報告を行い、6章で考察を延べ、7章で本稿のまとめと今後の課題を述べる。

2. 関連研究

2.1 アフィリエーションネットワーク分析

社会学の分野において、個人とコミュニティのつながりを分析するアフィリエーションネットワーク分析が提案されている²⁾。図1は個人とコミュニティの関係を表したアフィリエーションネットワークの一例である。丸が個人を、四角がコミュニティを表しており、個人が組織に属していることを実線で表す。図1のアフィリエーションネットワークモデルを元に、図2の個人と個人、図3のコミュニティ間のつながりを表す2つのネットワークモデルを用いて分析が行われる。

アフィリエーションネットワークモデルにおけるエッジは図1の場合は個人の所属情報を表し、図2は個人、図3はコミュニティを表す。このため、複数のコミュニティに属する個人間のコミュニケーション関係(例えばメールの送受信関係など)を分析する目的には用いることができない。

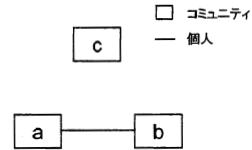


図3 コミュニティ間ネットワークモデル

3. OSS コミュニティにおけるコミュニケーションネットワークとコーディネータ

3.1 コミュニケーションネットワーク

OSS コミュニティはオンライン上の分散開発環境であるため、ML や掲示板を通じた非対面のコミュニケーションを基本とする。対面でのコミュニケーションが大きな役割を占める一般的なソフトウェア開発とは異なり、OSS コミュニティでは、ML や掲示板などのコミュニケーションメディアが参加者のコミュニケーションネットワークの形成に大きな影響を与える¹⁰⁾。したがって、OSS コミュニティにおけるコミュニケーションネットワークは、ML や掲示板などでのメッセージの送信者と返信者の関係から構成されるとみなすことができる^{1),10)}。

本稿では、分析対象とするコミュニティでの参加者同士のメッセージの送受信関係をエッジ(参加者同士のつながり)として OSS コミュニティのコミュニケーションネットワークを定義する。

3.2 コーディネータ

本稿では、開発者コミュニティとユーザコミュニティといった複数のサブコミュニティを調整し媒介する人物をコーディネータと呼ぶ。図4は OSS コミュニティの組織構造を表したものであり、黒丸はコーディネータを、白丸はそれ以外のコミュニティの参加者を表し、各参加者を結ぶ線はメッセージの送受信関係を表す。

本稿では、図4に示す参加者 a や b のように、複数のサブコミュニティと送受信関係のある参加者は何かしらの調整行動を図っていると考え、コーディネータとみなす。図4の参加者 a は、開発者、ユーザ、バグ報告者、それぞれのコミュニティの参加者とメッセージの送受信関係があり、開発者、ユーザ、バグ報告者、3つのコミュニティを結ぶコーディネータ(開発者&ユーザ&バグ報告者)である。参加者 b は、開発者コミュニティとユーザコミュニティの参加者それぞれとメッセージの送受信関係があり、開発者コミュニティとユーザコミュニティを結ぶコーディネータ(開発者&ユーザ)である。

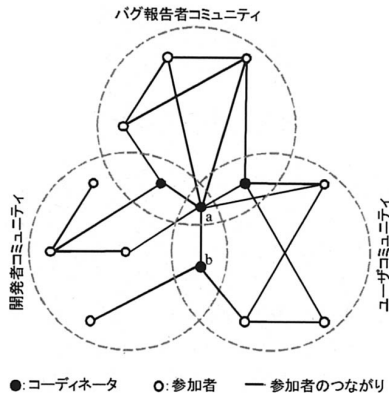


図 4 OSS コミュニティの組織構成

4. 分析方法

4.1 コミュニケーションネットワーク分析

メッセージの送信者と返信者の関係からコミュニケーションネットワークを構成し、コーディネータの総数といった統計量を用いて、そのネットワークを分析する。

本稿では、以下の統計量を用いる。

- サブコミュニティ間ごとのコーディネータの総数
- サブコミュニティごとの参加者の総数

4.2 時系列分析

コミュニケーションネットワークは時間の経過とともに変化するため、すべての期間のデータを用いてネットワークを求めることは不適切である。そのため、ある一定期間ごとにコミュニケーションネットワークを求める必要がある。

本稿では、ある一定期間 P ごとにコミュニケーションネットワークを求める際に、コミュニケーションネットワークを求める期間 P の開始時期を一定期間（例えば、 $P/2$ ）ずつスライドさせる Sliding Time Method を用いる⁵⁾。ある一定期間におけるコミュニケーションネットワークの求め方を図 5 に示す。ある一定期間のコミュニケーションネットワークはその期間の終了時期のものである。

本稿で用いる Sliding Time Method は、分析の期間が重複しない図 5 の (1) や (2) のような期間の区切り方より、コミュニケーションネットワークの変化を捉えやすい。例えば、図 5 に示すように $P_2 + P_3$ の期間で、Sliding Time Method によるコミュニケーションネットワークでは、コミュニケーションの活発さが反映されているが、(1)、(2) ではともに反映できてい

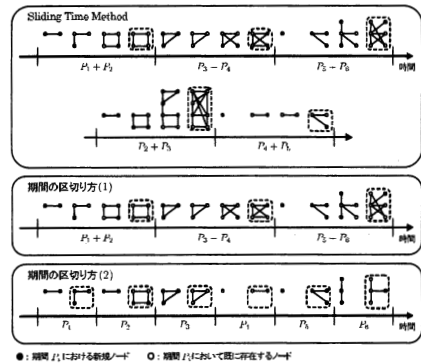


図 5 ネットワーク構造の時系列分析方法⁵⁾

ない。

本稿では、OSS コミュニティにおける開発者の社会的構造を分析した研究³⁾と同様、コミュニケーションネットワークを求める期間 P を3ヶ月、スライドさせる期間を1ヶ月とした。

5. ケーススタディ

本章では、PostgreSQLを対象としたケーススタディについて述べる。また、本稿では3つのコミュニティを対象とすることから、以下に挙げる4種類のコーディネータが存在すると考えられる。

- 開発者&ユーザ
- 開発者&バグ報告者
- ユーザ&バグ報告者
- 開発者&ユーザ&バグ報告者

ケーススタディでは、これら4種類のコーディネータと3つのサブコミュニティ（開発者、ユーザ、バグ報告者）それぞれの人数の推移について分析を行う。

5.1 PostgreSQL

PostgreSQL は BSD ライセンスに基づき配布されているフリーの関係データベース管理システム (RDBMS) である。商用/非商用を問わず無償で利用することができ、企業システムへの採用も増えていることから、商用の RDMS に劣らないソフトウェアとして広く認知されている。また、データベース管理システムであるという点から、特に高い信頼性が求められるソフトウェアであるといえる。

5.2 分析対象データ

分析対象データとしては各コミュニティの議論の場として主に用いられる ML を、各コミュニティの各 ML の利用者の集まりを単一コミュニティとしてとらえ、以下で各 ML について述べる。

- 開発者コミュニティ

表 1 各 ML の設置開始時期

pgsql-hackers	1996 年 1 月
pgsql-patches	1997 年 6 月
pgsql-general	1998 年 5 月
pgsql-bug	1996 年 12 月

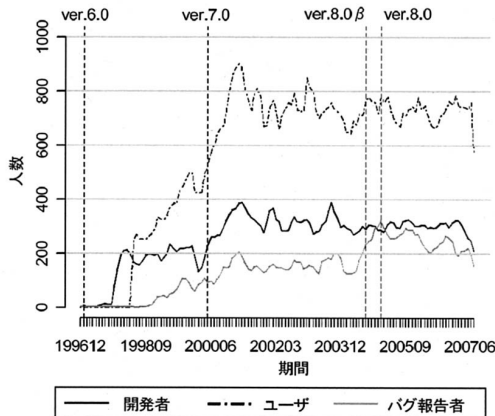


図 6 各コミュニティの総参加人数の推移

- postgresql-hackers : PostgreSQL の開発に関する話題を取り扱う ML であり、開発中のバグや新規機能についての議論が主として行われる。
- postgresql-patches : PostgreSQL の開発に関する話題を取り扱う ML であり、リリース後のバグの修正や新規機能についての議論が主として行われる。
- ユーザコミュニティ
 - postgresql-general : PostgreSQL のユーザが利用する ML であり、主として導入時のトラブルや SQL のバージョン問題など、開発者を交えて議論が行われる。
- バグ報告者コミュニティ
 - postgresql-bug : PostgreSQL のバグ報告専用の ML であり、主としてバグの修正についての議論が行われる。また、報告されたバグの修正が行われた場合は postgresql-patches へメッセージを送ることになっている。

分析期間としては、ML への投稿履歴が存在する全ての期間を利用したが、ML ごとにその設置時期が異なるため ML が存在しない期間については参加人数を 0 人として扱った。以下の表 1 に各 ML の設置時期を示す。

5.3 分析結果

各サブコミュニティに属する人数の推移を図 6 に、

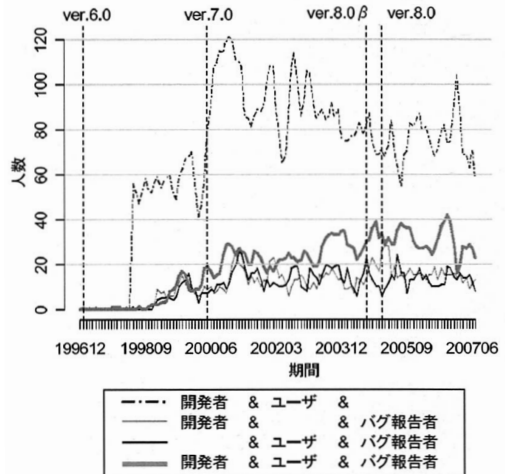


図 7 コーディネータの人数の推移

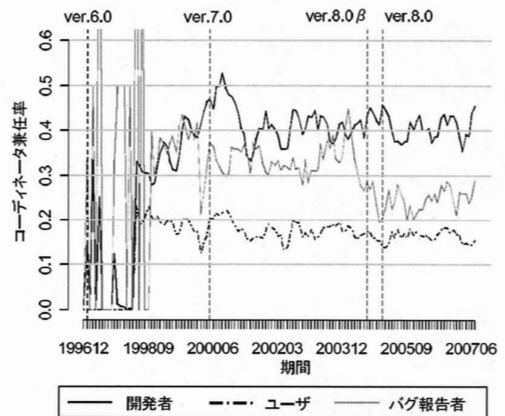


図 8 各コミュニティのコーディネータの割合の推移

コーディネータの人数の推移を図 7 に示す。また、コミュニティごとのコーディネータの割合を図 8 に示す。グラフの横軸は時間を表し、縦軸は、図 6 と図 7 は人数を、図 8 は割合を表す。また、縦の破線はリリース時期を示す。

各グラフについて、以下のような 3 つの特徴的な点を確認できる。

- (1) 図 7 と図 6 より、ver.7.0 のリリース前後にコーディネータ（開発者&ユーザ）の急激な増加（約 60 人から約 120 人）と同時にユーザ数が増加していることが確認できる。また、同時期に開発者コミュニティ内でのコーディネータの割合も増加していることがわかる（図 8）。
- (2) 図 7 より、ver.7.0 リリース以降から後期にかけ

て、コーディネータ（開発者&ユーザ）は減少傾向であるが、コーディネータ（開発者&ユーザ&バグ報告者）は緩やかに増加していることがわかる。

- (3) 図 8 より ver.8.0β リリースから ver.8.0 のリリースにかけてバグ報告者コミュニティ内のコーディネータの割合が急に減少している一方で、バグ報告者コミュニティ全体の参加人数（図 6）が増加していることが確認できる。

6. 考 察

- (1) ver.7.0 のリリース頃にコーディネータ（開発者&ユーザ）が増加している。これと同時期にユーザコミュニティの参加者が増加し、開発者コミュニティにおけるコーディネータの割合も増加していることがわかった。これは解釈の一つとして、ユーザコミュニティの参加者が増加し、ユーザからの要望などが増えたために多くの開発者がユーザのサポートをするためにユーザコミュニティにも参加し、コーディネータ（開発者&ユーザ）が増加したと考える。
- (2) 開発者コミュニティとユーザコミュニティを媒介するコーディネータはコーディネータ（開発者&ユーザ）だけでなく、コーディネータ（開発者&ユーザ&バグ報告者）も開発者コミュニティとユーザコミュニティを媒介している。分析結果で ver.7.0 をリリースして以降コーディネータ（開発者&ユーザ）は減少傾向にあり、コーディネータ（開発者&ユーザ&バグ報告者）は増加傾向にあることがわかった。また、図 1 に注目すると、開発者コミュニティ、ユーザコミュニティの参加者はほとんど変動していない。つまり開発者コミュニティとユーザコミュニティの参加者の合計人数にほとんど変動がないため、コーディネータ（開発者&ユーザ）の減少人数はコーディネータ（開発者&ユーザ&バグ報告者）へ立場が替わったと考えられる。これは一つの解釈として、バグ報告者に対して開発者が応えようとし、コーディネータ（開発者&ユーザ&バグ報告者）が 3 つのコミュニティの協調作業を行い始めていると考えられる。
- (3) ver.8.0 β, ver.8.0 のリリース頃においてバグ報告者コミュニティにおけるコーディネータの割合が減少し、同時期にバグ報告者コミュニティの参加人数が増加している。これから、バグ報告者コミュニティにおけるコーディネータの人

数が減少、もしくは変動せずに、新たにバグ報告者コミュニティに参加した個人が増えたと考えられる。人数の推移や割合を見ただけでは判断しかねるが、ひとつの解釈として、バグ報告者が増加している分、バグ報告が増加しているが、バグ報告者コミュニティにおけるコーディネータの人数が増加しないため、少人数のコーディネータがバグ報告者に対応していると考えられる。

7. おわりに

先行研究では開発者コミュニティ、ユーザコミュニティに注目し、コーディネータ（開発者&ユーザ）が重要であることが明らかとなった。本稿ではバグ報告者も重要な役割であると考え、バグ報告者コミュニティを含む OSS コミュニティのコミュニケーションネットワーク分析を行い、3 つのコミュニティにおけるコーディネータを分析した。ケーススタディとして扱った PostgreSQL において得られた知見を以下に示す。

- PostgreSQL においてもコーディネータ（開発者&ユーザ）が存在する。
- 3 つのコミュニティにおける 4 種類のコーディネータが確認できたが、コーディネータによってコーディネータの人数は大きく異なる。
- 協調作業を行うコーディネータの人数推移を分析することで、サブコミュニティとコーディネータとの関係を知ることができる。
- PostgreSQL では時間の経過に伴いコーディネータ（開発者&ユーザ）が減少し、コーディネータ（開発者&ユーザ&バグ報告者）が増加した。

3 つのサブコミュニティにおけるコーディネータについて、本稿では人数の推移のみで分析を行った。しかし、この分析方法では実際サブコミュニティ間でどれくらいのコミュニケーションをとっているかはわからない。また、個人がコーディネータとしての立場が替わる様子を分析することはできないため、今後は以下の 2 つを課題とする。

- 各サブコミュニティ、コーディネータの参加人数と個人間のコミュニケーションを同時に分析
- コーディネータの立場が替わる様子の分析

これらを行った上で、他の OSS コミュニティと比較し、他の OSS コミュニティとの違いを分析することを今後の目的とする。

謝辞 本研究の一部は、文部科学省「e-Society 基盤ソフトウェアの総合開発」の委託に基づいて行われたものである。また、本研究の一部は、文部科学省科

学研究補助費（基盤研究 B：課題番号 17300007，若手 B：課題番号 17700111）による助成を受けた。

参 考 文 献

- 1) Bird, C., Gourley, A., Devanbu, P., Gertz, M. and Swaminathan, A.: Mining Email Social Networks, *In Proceedings of the 2006 International Workshop on Mining Software Repositories (MSR'06)*, pp.137–143 (2006).
- 2) Everett, M.G. and Borgatti, S.P.: Calculating Role Similarities: An Algorithm that Helps Determine the Orbits of a Graph, *Social Networks* (1988).
- 3) Howison, J., Inoue, K. and Crowston, K.: Social Dynamics of Free and Open Source Team Communications, *In Proceedings of the 2nd International Conference on Open Source Systems (OSS'06)*, pp.319–330 (2006).
- 4) Jensen, C. and Scacchi, W.: Role Migration and Advancement Processes in OSSD Projects: A Comparative Case Study, *In Proceedings of the 29th International Conference on Software Engineering (ICSE'07)*, pp.364–374 (2007).
- 5) Kakimoto, T., Kamei, Y., Ohira, M. and Matsumoto, K.: Social Network Analysis on Communications for Knowledge Collaboration in OSS Communities, *In Proceedings of the International Workshop on Supporting Knowledge Collaboration in Software Development (KCS'D'06)*, pp.35–41 (2006).
- 6) Mockus, A., Fielding, R. T. and Herbsleb, J. D.: Two Case Studies of Open Source Software Development: Apache and Mozilla, *ACM Transactions on Software Engineering and Methodology*, Vol. 11, No. 3, pp. 309–346 (2002).
- 7) Nakakoji, K., Yamamoto, Y., Nishinaka, Y., Kishida, K. and Ye, Y.: Evolution Patterns of Open-Source Software Systems and Communities, *In Proceedings of the International Workshop on Principles of Software Evolution (IW-PSE'02)*, pp.76–85 (2002).
- 8) Netcraft Ltd.: Netcraft Web Server Survey, available from <http://www.netcraft.com/Survey/> (accessed 2007-06-30).
- 9) Raymond, E. S.: *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, O'Reilly and Associates (1999).
- 10) Yamauchi, Y., Yokozawa, M., Shinohara, T. and Ishida, T.: Collaboration with Lean Media: How Open-Source Software Succeeds, *In Proceedings of the 2000 Conference on Computer Supported Cooperative Work (CSCW'00)*, pp. 329–338 (2000).
- 11) Ye, Y. and Kishida, K.: Toward an Understanding of the Motivation Open Source Software Developers, *In Proceedings of the 25th International Conference on Software Engineering (ICSE'03)*, pp.419–429 (2003).
- 12) 大平雅雄, 裕本真佑, 前島弘敬, 亀井靖高, 松本健一: OSS コミュニティにおける共同作業プロセス理解のための中心性分析, グループウェアとネットワークサービスワークショップ2007, p.(to appear) (2007).