

コミュニティ媒介性： OSS 開発における協調作業への媒介度を評価する指標

大平雅雄¹ 杉本真佑¹ 松本健一¹

1. はじめに

一般に、Linux や Apache をはじめとする Open Source Software (OSS) 開発コミュニティには、情報共有や作業の効率化を図るために、参加者の役割(開発者、バグ報告者、ユーザなど)に応じたいくつかのサブコミュニティが存在する。参加者が地理的に分散している環境下において高機能かつ高品質な OSS を持続的に開発するためには、それぞれのサブコミュニティで行われる活動の結果として得られる成果物(ソースコード、バグ報告、ユーザフィードバックなど)を有機的に機能させ OSS に反映していくことが重要となる。

OSS コミュニティ参加者の内、コアメンバはいくつかのサブコミュニティに参加しコミュニティ全体の秩序の形成や組織化に貢献し、サブコミュニティ間の協調作業を円滑化する役割を担っている。我々の先行研究は、複数のサブコミュニティを媒介するコアメンバのコミュニケーション構造(図 1)が、成功コミュニティと衰退コミュニティとは大きく異なることを示した(図 2) [1]。また、異なるサブコミュニティの参加者をコアメンバが媒介する度合いを評価する指標として、従来の媒介中心性は適切ではないことも明らかとなった。本稿では、サブコミュニティ間の協調作業を円滑化するコアメンバの媒介度合いを定量的に評価するためのコミュニティ媒介性指標を紹介する。

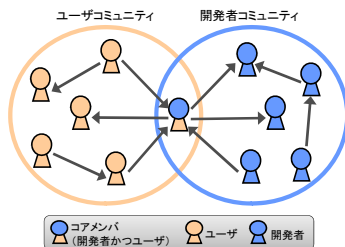


図 1 OSS コミュニティにおけるコミュニケーション構造の一例(ユーザコミュニティと開発者コミュニティ)

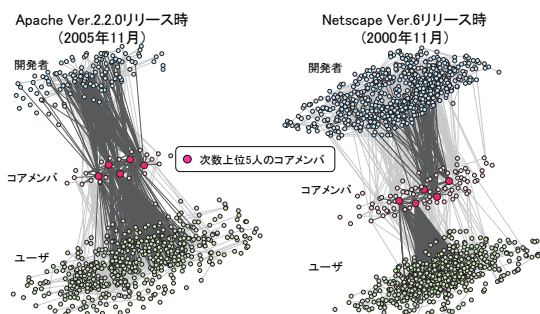


図 2 Apache コミュニティと
衰退期の Netscape コミュニティとの比較

2. コミュニティ媒介性

本研究では、コミュニティ参加者の ML への投稿と投稿に対する返信との関係から、参加者同士の繋がり(コミュニケーション構造)を定義する。図 1 は OSS コミュニティの参加者のコミュニケーション構造を表したものである。コミュニティ参加者をノードで表し、ML への投稿に対する返信をエッジで表している。

媒介中心性とは、評価対象のノードが他のノードに対して媒介者としての役割をどの程度果たしているかを示す指標である。評価対象のノードを介して情報伝達可能なノード対の総数によって評価する。先行研究では、コアメンバの媒介の度合いを評価するために、媒介中心性を用いた [1]。しかし、複数のサブコミュニティが存在する OSS コミュニティにおけるコアメンバの媒介度合いを評価するための指標として、媒介中心性は以下のような問題があることがわかった。

- (1) コミュニティ内/間の媒介を区別しない。
- (2) 情報伝達の向きを区別しない。

特に本研究では、サブコミュニティ間にまたがる情報伝達におけるコアメンバの媒介度合いの評価を対象としている。そのため、従来の媒介中心性では、コミュニティ内の媒介への偏りが顕著なノード(たとえば図 2 の Netscape のコアメンバ)であっても、サブコミュニティ間を媒介するコアメンバとして高く評価される場合があり、本研究の目的とは合致しない指標であった。

上述の媒介中心性の問題点に対して、コアメンバがコミュニティ間を媒介する構造に着目することで問題の解決を図る。コアメンバがコミュニティ間を媒介する構造は、一方のコミュニティからコアメンバを介してもう一方のコミュニティへつながるエッジの組み合わせによって表せる。そのため、コアメンバによる媒介の流れには、サブコミュニティ X からサブコミュニティ Y への媒介と、サブコミュニティ Y からサブコミュニティ X への媒介、2 つの流れが考えられる。例えば、開発者とユーザの協調作業を支えるコアメンバを特定するためには、コアメンバが開発者コミュニティとユーザコミュニティを相互に媒介する度合いを評価する必要がある。そのため本研究では、サブコミュニティ X からサブコミュニティ Y を媒介する度合いとサブコミュニティ Y からサブコミュニティ X を媒介する度合いの調和平均に基づいて、コアメンバがサブコミュニティ間を相互に媒介する度合いを評価するコミュニティ媒介性を構築した。コミュニティ媒介性は、OSS コミュニティにおけるコアメンバの媒介のパターンを想定した 3 つの指標からなる。

本研究で想定している媒介の 3 つパターンを図 3 (簡単のために片方向の媒介のみ)に示す。図 3 のエッジの太さは、各ノードとコアメンバ (v_i) との間で行われた ML へのメッセージ投稿数に比例したものであり、ノードの大きさは、各ノードが有するエッジの数(次数)に比例したものを表している。

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

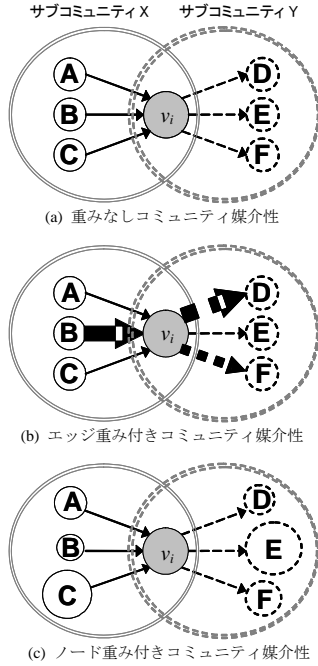


図3 媒介のパターン

重みなしコミュニティ媒介性 (C_{ICB})

C_{ICB} は、媒介の方向のみに着目した指標である。各ノードを投稿数や回数に関わりなく同等に扱うため、媒介のパターンは図3 (a) のように扱われる。

$$C_{ICB}(v_i) = \frac{2f_x(v_i)f_y(v_i)}{f_x(v_i) + f_y(v_i)} \quad (1)$$

$$f_x(v_i) = \frac{InDegree_x(v_i)}{N_x - 1}, \quad f_y(v_i) = \frac{OutDegree_y(v_i)}{N_y - 1}$$

$f_x(v_i)$: コミュニティ X のノード数に対する v_i の入次数の割合
$f_y(v_i)$: コミュニティ Y のノード数に対する v_i の入次数の割合
$InDegree_x(v_i)$: コミュニティ X から v_i への次数
$OutDegree_y(v_i)$: v_i からコミュニティ Y への次数
N_x	: コミュニティ X のノード数
N_y	: コミュニティ Y のノード数

エッジ重み付きコミュニティ媒介性 ($C_{ICBwithWE}$)

$C_{ICBwithWE}$ は、ML へのメッセージ投稿数に着目した指標である。各ノードとコアメンバとの間で行われた ML への投稿数が多いほどコアメンバが情報伝達に寄与しているとして、コアメンバの媒介度合いの評価に反映する。媒介のパターンは図3 (b) のように扱われる。ノード B からノード D への情報伝達のように、送信回数の多いノード間を媒介するほど媒介度合いが相乗的に高まると考える。

$$C_{ICBwithWE}(v_i) = \frac{2f_x(v_i)f_y(v_i)}{f_x(v_i) + f_y(v_i)} \quad (2)$$

$$f_x(v_i) = \frac{InEdgeWgt_x(v_i)}{N_x - 1}, \quad f_y(v_i) = \frac{OutEdgeWgt_y(v_i)}{N_y - 1}$$

$f_x(v_i)$: v_i からコミュニティ X へのエッジ重みの総和
$f_y(v_i)$: v_i からコミュニティ Y へのエッジ重みの総和
$InEdgeWgt_x(v_i)$: コミュニティ X から v_i への総メッセージ送信数
$OutEdgeWgt_y(v_i)$: v_i からコミュニティ Y への総メッセージ送信数
N_x	: コミュニティ X のノード数
N_y	: コミュニティ Y のノード数

ノード重み付きコミュニティ媒介性 ($C_{ICBwithWN}$)

$C_{ICBwithWN}$ は、各ノードの次数 (ノードの重み) に着目した指標である。次数が多いノードほどコミュニティ内の情報について熟知し、また得られた情報をコミュニティ内に広めやすいという考えに基づき、コア

メンバの媒介度合いの評価に反映される。媒介のパターンは図3 (c) のように扱われる。ノード C からノード E への情報伝達のように、次数の多いノード間を媒介するほど媒介の度合いが相乗的に高まると考える。

$$C_{ICBwithWE}(v_i) = \frac{2f_x(v_i)f_y(v_i)}{f_x(v_i) + f_y(v_i)} \quad (3)$$

$$f_x(v_i) = \frac{InNodeWgt_x(v_i)}{N_x - 1}, \quad f_y(v_i) = \frac{OutNodeWgt_y(v_i)}{N_y - 1}$$

$f_x(v_i)$: v_i からコミュニティ X に対するノード重みの総和
$f_y(v_i)$: v_i からコミュニティ Y に対するノード重みの総和
$InNodeWgt_x(v_i)$: v_i とつながるコミュニティ X のノードの次数の総和
$OutNodeWgt_y(v_i)$: v_i とつながるコミュニティ Y のノードの次数の総和
N_x	: コミュニティ X のノード数
N_y	: コミュニティ Y のノード数

3. 議論

提案指標の有効性を確認するために、Apache コミュニティを対象としたケーススタディを行った[2]。その結果、従来指標である媒介中心性と比較して、コミュニティ媒介性は 3 種類の媒介のパターンをより適切に評価できることを確認した。

コミュニティ媒介性はコミュニティの規模を考慮した指標であるため、コミュニティの規模に変化があった場合 (成長や衰退など) でも、同一の指標を用いることができる。そのため、コアメンバの媒介度合いを異なる時期で比較し、コミュニティの運営方針や協調作業円滑化のための対策を立てる際のデータとして役立つものと期待される。また、コミュニティ媒介性は、OSS コミュニティにおけるコアメンバの媒介度合いの評価に限らず、近年盛んに行われているオフショア開発のドメインで応用できると考えられる。分散開発の形態をとるオフショア開発は、発注先と発注元のコミュニケーション及び協調作業を仲介するブリッジ SE の能力が鍵を握る。ブリッジ SE の能力を定量的に測ることができれば、プロジェクトの規模や複雑さに対応できる能力を有する適切なブリッジ SE をプロジェクトに配置することが可能になる。

4. まとめと今後の課題

本稿では、OSS 開発コミュニティのコアメンバが、複数のサブコミュニティを相互に媒介する度合いを評価するための指標であるコミュニティ媒介性を紹介した。本研究の今後の課題は、コアメンバ同士の繋がりを考慮して指標を改良すること、オフショア開発など応用ドメインへ適用し有効性を確認することなどが挙げられる。

謝辞

本研究を遂行するにあたって、日本電気 (株) の前島弘敬氏に多くの御協力と御指導を頂いた。深く感謝いたします。本研究の一部は、文部科学省「次世代 IT 基盤構築のための研究開発」の委託に基づいて行われた。また、本研究の一部は、文部科学省科学研究補助費 (若手 B: 課題番号 20700028) および公益信託マイクロソフト知的財産研究助成基金による助成を受けた。

参考文献

- [1] 大平雅雄, 松本真佑, 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, 情報処理学会シンポジウムシリーズ, Vol.2007, No.11, pp.7-12, 2007.
- [2] 前島弘敬, コミュニティ媒介性: OSS コミュニティにおけるコアメンバの媒介度合いを評価するための指標, Master's Thesis, 奈良先端科学技術大学院大学 情報科学研究科, 2008.