

コーディネータのコミュニティ媒介性の評価指標の提案

前島弘敬[†] 梶本真佑[†] 亀井靖高[†] 柿元健[†] 大西洋司[†] 大平雅雄[†] 松本健一[†]

Open Source Software (OSS) 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す。我々は実際の OSS コミュニティにおけるコーディネータの仲介の度合いを、Freeman の中心性の1つである媒介中心性を用いて評価し、媒介中心性ではコーディネータの仲介の度合いを十分に表現できないことを明らかにした。本稿では、コーディネータが2つのコミュニティを仲介する度合いを仲介フローに基づいて評価するコミュニティ媒介性を提案する。9通りのネットワークに対して提案指標を適用した評価実験の結果、提案指標は媒介中心性と比較して、コーディネータによるコミュニティ間の情報伝達の偏りを考慮に入れて仲介の度合いを表現することができた。

Inter-Community Betweenness: Metrics for Measuring Coordination Quality in OSS Development

HIROTAKA MAESHIMA[†] SHINSUKE MATSUMOTO[†] YASUTAKA KAMEI[†]
TAKESHI KAKIMOTO[†] YOJI ONISHI[†] MASAO OHIRA[†] KEN-ICHI MATSUMOTO[†]

Since the Open Source Software (OSS) community depends on collaboration among participants such as developers and end-users, coordinators play important roles of information transmission and adjustment between developers and users. Our previous study indicated that Freeman's *betweenness centrality* cannot sufficiently represent intermediacy skills of coordinators in OSS communities. In this paper, we propose *Inter-Community Betweenness* which is metrics for assessing the degree of coordinators' intermediacy based on a flow of intermediacy. As the experiment results of applying the proposed metrics to 9 social networks, we have confirmed that *Inter-Community Betweenness* can represent the degree of coordinators' intermediacy in consideration of coordinators' inequable information transmission between communities.

1. はじめに

近年、Apache や Mozilla に代表される Open Source Software (OSS) が注目を浴びている[1][5]。OSS とは、ソフトウェアのソースコードが無償で公開されており、改良、再配布が可能なソフトウェアのことである。OSS は、商用ソフトウェアに引けを取らない品質・性能を備えており、安価で利用可能である。従って、多くのユーザに支持されており、企業システムの一部に導入されている事例も多く、OSS は社会的基盤として広く利用されている有用なソフトウェアと言える。

一方で、大多数の OSS はユーザからの支持が受けられず、開発が衰退・中断している現状が指摘されている[7]。OSS 開発の衰退・中断は、ユーザに対するサポートも同時に打ち切られるため、個人や企業の OSS 利用を妨げる要因となっている。そのため、OSS 開発の実態を分析し、OSS 開発を成功に導く要因を明らかにすることで、OSS 開発を支援する研究が行われている[6]。

OSS 開発のコミュニティ (OSS コミュニティ) はユーザの存在が重要であり、ユーザ参加型コミュニティで

あると言われている [8]。ユーザからの機能追加要求やバグ報告、意見を開発者に重要な情報として伝達し、伝達された情報を開発者が OSS 開発にフィードバックすることで、OSS の更なる改良へと繋がる[4][10]。そのため、ユーザを共同開発者と位置付けることもある[8]。

しかし、開発者、ユーザそれぞれの意見は必ずしもうまく相互に伝達されているとは限らない。一般に、OSS コミュニティには、メーリングリスト (ML) や掲示板などのコミュニケーション媒体が目的や役割に応じて複数設置されており、開発者とユーザは通常異なる ML や掲示板で議論している。そのため、OSS コミュニティ内に、開発者とユーザのサブコミュニティ (開発者コミュニティとユーザコミュニティ) が個別に形成されており、両者の情報伝達を阻害する要因となっている (以降、「OSS コミュニティ」を除いて「コミュニティ」は OSS コミュニティ内の開発者やユーザによるサブコミュニティを指すこととする)。従って、ユーザコミュニティ内で議論された開発に有益な情報が、開発者コミュニティの参加者にまで伝達されない可能性がある。

開発者コミュニティとユーザコミュニティの情報伝達には、開発者とユーザの両方のコミュニティに属し、参加者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す[3]。近年、コーディネータに注

[†]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

目し、コーディネータと OSS 開発の成功の関係に着目した研究も行われている [1][5].

我々はこれまで、3つの OSS コミュニティにおいてソーシャルネットワーク分析[9]の分野で提案されている Freeman の中心性[2]の1つである媒介中心性を用いてコミュニティ参加者の情報伝達の仲介の度合いの評価を行ない、活動が活発な OSS コミュニティには多くの仲介を行っているコーディネータと考えられる人物が存在することを確認した[11]. しかし、媒介中心性はコーディネータを評価する指標として、不十分であることも明らかとなった。(2.2節参照)

本稿では、先行研究で明らかとされた媒介中心性の問題点を改善し、ソーシャルネットワーク分析に基づいたコーディネータがコミュニティ間の仲介を行う度合いを評価するための指標を提案する。

以降、2章で従来指標である媒介中心性とその問題点について述べ、3章でコーディネータによるコミュニティ間の仲介の度合いを評価する指標を提案する。4章で9通りのネットワークに対して、提案指標を適用した評価実験とその結果について報告する。5章で、本稿のまとめを述べる。

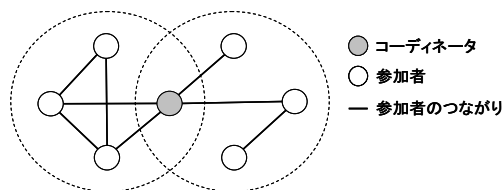
2. 媒介中心性

本稿では、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仲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、コーディネータの仲介の度合いを評価するために、媒介中心性を用いた。しかし、媒介中心性による評価では、以下の点において不十分であると考えられる。

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

図2は、コーディネータの仲介の種類を表している。ここでは、コーディネータの仲介には、コミュニティ内とコミュニティ間の2種類あることが分かる。しかし、本稿で対象とするコーディネータの評価は、コミュニティ間にまたがる情報伝達である。そのため、媒介中心性では、コミュニティ内の仲介に偏りが顕著なノードであっても、コーディネータ間を仲介するコーディネータとして評価する可能性がある。



開発者コミュニティ ユーザコミュニティ

図1 OSS コミュニティのコミュニケーション構造

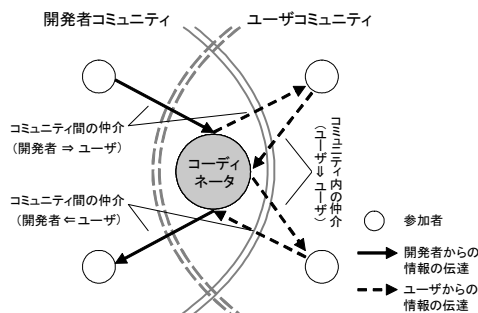


図2 コーディネータの仲介の種類

また、媒介中心性はノード対の数によって仲介の度合いを評価するため、情報伝達の向きを区別しない。たとえば、ノード a からノード b と、ノード b からノード a の異なる向きの仲介を考えた場合、ノード対という考え方から双方は同一視される。しかし、情報伝達の向きを区別しないと、開発者からユーザへの仲介とユーザから開発者への仲介の度合いが同じ、両者の違いを判別することができない。

そこで本稿では、コーディネータの仲介の度合いを評価するために、コーディネータのコミュニティ間の情報伝達だけに着目し、また、情報伝達の方向を区別してコーディネータの仲介の度合いを評価するコミュニティ媒介性を提案する。

3. コミュニティ媒介性

3.1 キーアイデア

本稿では、コーディネータの仲介の度合いを、コーディネータが情報伝達可能な各コミュニティのノードとエッジの数に基づき評価する。エッジは各ノードをつなぐ線であり、情報伝達の向きを表す。

コーディネータの仲介の度合いを評価するために、図3に示すような、一方のコミュニティからコーディネータ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着目し、これらを仲介フローと呼ぶ。情報伝達される向きが異なる場合を別に扱うことで、それぞれのコミュニティの視点からコーディネータの仲介の度合いを評価することができる。

3.2 コミュニティ媒介性

本節では、コーディネータの情報伝達を表す仲介フローに基づき3つのコミュニティ媒介性を提案する。

図4に本稿で考慮する仲介フローの3つパターンを示す。図4のエッジの太さは、各ノードとコーディネータとの間で行われたMLへの投稿数を示し、ノードの大きさは、各ノードが有するエッジの数(次数)を示す。

重みなしコミュニティ媒介性

1つ目の指標は、仲介フローの有無のみに着目した指標である。重みなしコミュニティ媒介性では、各ノードを投稿数や次数に関わりなく同等に扱うため、仲介フローは図4(a)のように扱われる。重みなしコミュニティ媒介性は以下の式で算出される。

$$C_{ICB} = \frac{InDegree_x}{N_x} \times \frac{OutDegree_y}{N_y} \quad (1)$$

- C_{ICB} : 重みなしコミュニティ媒介性
- $InDegree_x$: コミュニティXからコーディネータへの次数
- $OutDegree_y$: コーディネータからコミュニティ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

2つ目の指標は、仲介フローにおけるMLへの投稿数に着目した指標である。エッジ重み付きコミュニティ媒介性では、各ノードとコーディネータとの間で行われたMLへの投稿数に着目し、投稿数が多いほどコーディネータが情報伝達に寄与しているとして、コーディネータの仲介の度合いの評価に反映されている。エッジの重みを考慮した仲介フローは図4(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、投稿数の多いノード間を仲介するほどコーディネータの仲介の度合いが相乗的に高まるものとする。エッジ重み付きコミュニティ媒介性は以下の式で算出される。

$$C_{ICBwithWE} = InEdgeWgt_x \times OutEdgeWgt_y \quad (2)$$

$$InEdgeWgt_x = \sum_{i=1}^{InDegree_x} EdgeWgt_{xi} \quad (3)$$

$$OutEdgeWgt_y = \sum_{j=1}^{OutDegree_y} EdgeWgt_{yj} \quad (4)$$

- $C_{ICBwithWE}$: エッジ重み付きコミュニティ媒介性
- $EdgeWgt_{xi}$: コミュニティXのノードiからコーディネータへの投稿数
- $EdgeWgt_{yj}$: コーディネータからコミュニティYのノードjへの投稿数

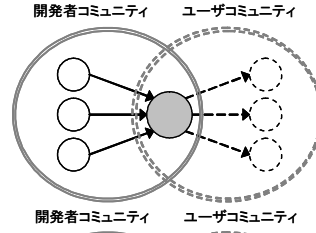
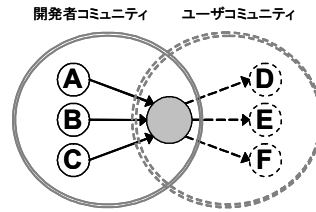
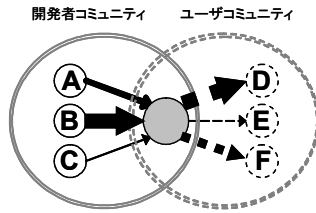


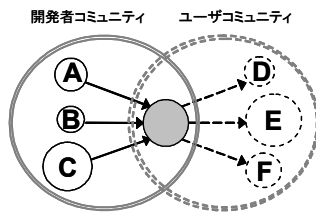
図3 仲介フローの例



(a) 重みなしコミュニティ媒介性



(b) エッジ重み付きコミュニティ媒介性



(c) ノード重み付きコミュニティ媒介性

図4 仲介フローのパターン

ノード重み付きコミュニティ媒介性

3つ目の指標は、仲介フローにおける各ノードの次数(ノードの重み)に着目した指標である。ノード重み付きコミュニティ媒介性では、各ノードの次数に着目する。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、コーディネータの仲介の度合いに反映されている。ノードの重みを考慮した仲介フローは図4(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仲介するほ

どコーディネータの仲介の度合いが相乗的に高まるものとする。ノード重み付きコミュニティ媒介性は以下の式で算出される。

$$C_{ICwithWN} = \frac{InNodeWgt_x}{N_x} \times \frac{OutNodeWgt_y}{N_y} \quad (5)$$

$$InNodeWgt_x = \sum_{i=1}^{InDegree_x} NodeWgt_{xi} \quad (6)$$

$$OutNodeWgt_y = \sum_{i=1}^{OutDegree_y} NodeWgt_{yi} \quad (7)$$

$C_{ICwithWN}$: ノード重み付きコミュニティ媒介性
 $NodeWgt_{xi}$: コーディネータとつながっているコミュニティ X のノード i の次数
 $NodeWgt_{yj}$: コーディネータとつながっているコミュニティ Y のノード j の次数

4. 評価実験

4.1 実験設定

本稿では、提案する 3 つのコミュニティ媒介性が実際の OSS コミュニティでも適用可能であることを確認す
 表 1 実験に用いるネットワークの種類とるために評価実験を行う。そのために、以下に示す OSS コミュニティにおけるコミュニケーション構造の特徴に基づき、

評価対象とする指標

	変化させるパラメータ	評価対象とする提案指標
$\alpha 1 \sim \alpha 3$	次数	重みなしコミュニティ媒介性
$\beta 1 \sim \beta 3$	エッジの重み	エッジ重み付きコミュニティ媒介性
$\gamma 1 \sim \gamma 3$	ノードの重み	ノード重み付きコミュニティ媒介性

評価対象となるネットワークを構築する。

- OSS コミュニティ参加者同士の繋がりの強さを表す投稿数 (エッジの重み) は参加者ごとに異なる。
- OSS コミュニティ参加者のコミュニティ内での活動の程度を表す次数 (ノードの重み) は参加者ごとに異なる。

本実験では、コーディネータ自身のノードの重みである次数と、コーディネータとつながるノードのエッジの重み、また、ノードの重みの 3 点に着目し、それぞれの違いをパラメータ (次数、エッジの重み、ノードの重み) として 3 段階に変化させた。

図 5 に実験に用いる 9 通りのネットワークを、表 1 に各ネットワークで変化させるパラメータと評価の対象となるコミュニティ媒介性を示す。図 5 では、各ネットワークの中央に位置するノードをコーディネータとして示し、コーディネータを境界として上をコミュニティ

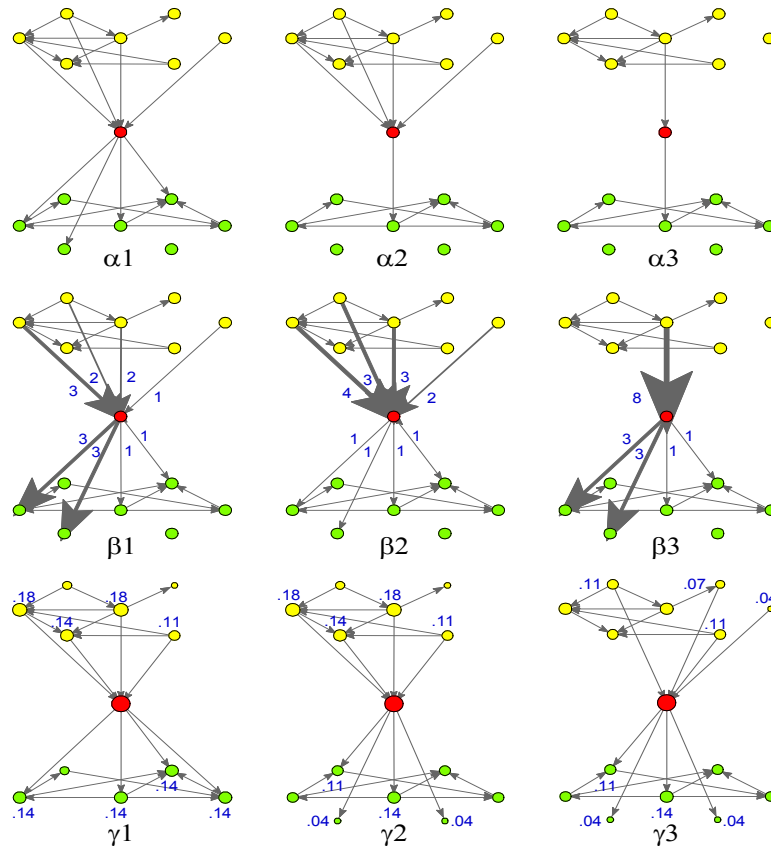


図 5 実験に用いる 9 通りのネットワーク

X, 下をコミュニティ Y として, 別のコミュニティに属するノードが配置されている。また, エッジの太さはエッジの重みを, ノードの大きさはノードの重みを表す。

9通りのネットワークでは, 各パラメータ (次数, エッジ重み, ノード重み) を変化させ, コミュニティ X, Y ともに各パラメータの値が高い $\alpha 1, \beta 1, \gamma 1$ と, コミュニティ X のみの各パラメータが高い $\alpha 2, \beta 2, \gamma 2$ に加え, コミュニティ X, Y ともに各パラメータの低い $\alpha 3, \beta 3, \gamma 3$ を設けている。

また, 提案する 3 つのコミュニティ媒介性によって, コーディネータの仲介の度合いが適切に評価できていることを確認するために, 比較対象として媒介中心性を用いる。

コミュニティ媒介性では, コーディネータのコミュニティ X からコミュニティ Y への情報伝達と, Y から X への情報伝達から, コーディネータの仲介の度合いを評価することができるが, 紙面の都合上, 本実験では一方の情報伝達の向き (コミュニティ X からコミュニティ Y) のみを扱う。

4.2 結果と考察

表 2 に 9通りのネットワークにおけるコミュニティ媒介性と媒介中心性の値を示す。ネットワーク $\alpha 1 \sim \alpha 3$ において重みなしコミュニティ媒介性と媒介中心性を, ネットワーク $\beta 1 \sim \beta 3$ においてエッジ重み付きコミュニティ媒介性と媒介中心性を, ネットワーク $\gamma 1 \sim \gamma 3$ においてノード重み付きコミュニティ媒介性と媒介中心性を, それぞれ比較する。

重みなしコミュニティ媒介性と媒介中心性との比較

コーディネータの入次数および出次数が異なるネットワーク $\alpha 1 \sim \alpha 3$ において, コーディネータの入次数と出次数の総和が大きくなるほど, 重みなしコミュニティ媒介性, 媒介中心性とも値は増加している。また, $\alpha 1 \sim 3$ で比較すると $\alpha 3$ を 1 とした時の比率は, 重みなしコミュニティ媒介性は約 (17:4:1), 媒介中心性は約 (3:3:1) となっている。従って, 媒介中心性と比べて重みなしコミュニティ媒介性は次数の変化を顕著に表しているといえる。

また, 重みなしコミュニティ媒介性はそれぞれのコミュニティのノード数の値で正規化した値を乗算した指標であるので, $\alpha 1$ のようなそれぞれのコミュニティとも 7人中 4人とコミュニティに属する多くの人がコーディネータとのエッジを持っている点も値に表れている

といえる。

エッジ重み付きコミュニティ媒介性と媒介中心性との比較

コーディネータへの投稿数が均一な $\beta 1$ と投稿数に偏りを持つ $\beta 2$ のネットワークで比較すると, エッジ重み付きコミュニティ媒介性の値は, コーディネータとの投稿数が多いほど増加している。一方, 媒介中心性では, 投稿数を考慮しないため, $\beta 1$ と $\beta 2$ のように重みを考慮しない仲介フローの構造が同じ場合には値に差が生じていない。

また, コミュニティ X からの総投稿数は等しいが, 入次数が異なる $\beta 1$ と $\beta 3$ のネットワークを比較すると, エッジ重み付きコミュニティ媒介性は同じ値を取っているが, 他の指標では値が低下している。

ノード重み付きコミュニティ媒介性と媒介中心性との比較

コーディネータとつながっているノードの重みが異なるネットワーク $\gamma 1 \sim \gamma 3$ において, コーディネータとつながっているノードの重みが大きいほど, ノード重み付きコミュニティ媒介性の値は大きくなっている。一方で, 媒介中心性はノードの重みが大きい $\gamma 1$ では最も値が小さくなっている。

これらの結果は, ノード重み付きコミュニティ媒介性がコミュニティに属する中心的なノード同士を媒介している度合いを表すのに対して, 媒介中心性がコミュニティに属する全ノードを網羅的に媒介する際に高い値をとる指標であるためだと考えられる。

コミュニティ媒介性と媒介中心性との比較

3種類のコミュニティ媒介性と媒介中心性を比較すると, 仲介フローのパラメータの変化による媒介中心性の値の変化は, コミュニティ媒介性の値の変化に対して相対的に小さくなっている。これは, 媒介中心性ではコミュニティ内の仲介を考慮しているにも関わらず, 今回の評価実験では, コミュニティ間の仲介の視点からのみ見ているためであると考えられる。すなわち, 媒介中心性はコミュニティ内の仲介を考慮するため, コミュニティ間の仲介を行うコーディネータの仲介の度合いのみを表現するには不十分であるといえる。

本実験で得られた知見を以下にまとめる。

- 重みなしコミュニティ媒介性は媒介中心性と比較して, コーディネータの各コミュニティとの次数の偏りの大きさ (例えば, $InDegreeX = 1,$

表 2 コーディネータの仲介の度合いの評価結果

	$\alpha 1$	$\alpha 2$	$\alpha 3$	$\beta 1$	$\beta 2$	$\beta 3$	$\gamma 1$	$\gamma 2$	$\gamma 3$
重みなしコミュニティ媒介性	0.33	0.08	0.02	0.33	0.33	0.08	0.33	0.33	0.33
エッジ重み付きコミュニティ媒介性	16.00	4.00	1.00	64.00	48.00	64.00	16.00	16.00	16.00
ノード重み付きコミュニティ媒介性	0.86	0.21	0.08	0.86	0.86	0.31	1.30	0.69	0.37
媒介中心性	0.16	0.14	0.05	0.16	0.16	0.07	0.14	0.19	0.19

$OutDegreeY=10$) を顕著に表現することが可能である。

- エッジ重み付きコミュニティ媒介性は媒介中心性と比較して、コーディネータの各コミュニティへの総投稿数の偏りの大きさ（例えば、 $InEdgeWgtX=1$, $OutEdgeWgtY=10$) を顕著に表現することが可能である。
- 媒介中心性はコミュニティに属する全ノードを網羅的に媒介している度合いを表現するのに対して、ノード重み付きコミュニティ媒介性はコミュニティに属する中心的なノード同士を媒介している度合いを表現することが可能である。

5. おわりに

本稿では、コーディネータによる開発者コミュニティとユーザコミュニティの仲介の度合いを評価するために、コーディネータの仲介フローから仲介の度合いを求める3つのコミュニティ媒介性を提案した。9通りのネットワークを用いた評価実験から、従来の媒介中心性と比較して3つの提案指標の有意性を確認することができた。

また、本稿では、3つのコミュニティ媒介性を提案しており、3つの指標によるコーディネータの多面的な評価によって、OSSコミュニティの活動状態を分析・把握することが可能であると考えられる。たとえば、時系列による値の推移に着目したとき、重みなしコミュニティ媒介性の値の増加に対して、エッジ重み付きコミュニティ媒介性の値に変化が乏しい場合、それぞれのコミュニティに対しての繋がりは増える一方で、個々との繋がりが希薄になっている傾向にあると考えることができる。また、ノード重み付きコミュニティ媒介性の増加に対して、重みなしコミュニティ媒介性に変化が乏しい場合、コーディネータ自身ではなく、コミュニティ内の活動が活性化している傾向にあると考えることができる。

今後の課題として、OSSコミュニティのMLや掲示板などコミュニケーション履歴を用いて提案指標を適用し、実際に情報伝達や調整行動を行うコーディネータを抽出できるかどうかを調べる必要がある。

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) Freeman, L.C.: Centrality in Social Networks: Conceptual Clarification, Social Networks, Vol.1, No.3, pp.215-239 (1979).
- 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) 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).
- 6) 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 (IWPSE'02), pp.76-85 (2002).
- 7) Ohira, M., Ohoka, T., Kakimoto, T., Ohsugi, N. and Matsumoto, K.: Supporting Knowledge Collaboration Using Social Networks in A Large Scale Online Community of Software Development Projects, In Proceedings of the International Workshop on Supporting Knowledge Collaboration in Software Development (KCS'D'05), pp.835-840 (2005).
- 8) Raymond, E.S.: The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary, O'Reilly and Associates (1999).
- 9) Wasserman, S. and Faust, K.: Social Network Analysis: Methods and Applications, Cambridge University Press (1994).
- 10) 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).
- 11) 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, 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 (2007) (to appear).