

## ユーザのコスト多様性を考慮した報酬配分方法の評価

橋本 遼<sup>†</sup> 佐藤健一朗<sup>†</sup> 新熊 亮一<sup>††</sup> 高橋 達郎<sup>††</sup>

<sup>††</sup> 京都大学大学院 情報学研究科  
〒 606-8501 京都市左京区吉田本町

E-mail: †{hashimoto,sato}@cube.kuee.kyoto-u.ac.jp, ††{shinkuma,ttakahashi}@i.kyoto-u.ac.jp

あらまし P2P 情報共有では、コンテンツ提供など参加ユーザの貢献が不可欠である。しかし、貢献行動はリソースの拠出を伴うため、心理的な負担であるコストが発生する。このため、貢献行動を行わないフリーライダーと呼ばれるユーザが多くを占めることが問題となっている。そこで、ユーザの貢献に適切な報酬を与えて貢献を促進するインセンティブメカニズムが考案されており、貢献行動に伴う主観的なコストに対応可能なインセンティブメカニズムとして EMOTIVER が提案されている。また、コストを直接的コストと間接的成本に分類したときに、支配的なコストと有効な貢献評価関数の関係が明らかにされている。本稿では、ユーザのコストがより多様な環境で、直接的・間接的成本の支配性がユーザの貢献度を与える影響についてシミュレーション評価を行い、コストに対する EMOTIVER のロバスト性を確認する。また、貢献評価関数の最適性を評価するため、理想的な報酬配分が行われたときの結果とも比較を行う。そして、理想値との差を補償する新たな報酬配分方法を考案する。

キーワード P2P 情報共有, インセンティブ報酬, コスト

## Evaluation of Rewarding Methods Considering Users' Cost Variety

Ryo HASHIMOTO<sup>†</sup>, Kenichiro SATO<sup>†</sup>, Ryoichi SHINKUMA<sup>††</sup>, and Tatsuro TAKAHASHI<sup>††</sup>

<sup>††</sup> Graduate School of Informatics, Kyoto University  
Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

E-mail: †{hashimoto,sato}@cube.kuee.kyoto-u.ac.jp, ††{shinkuma,ttakahashi}@i.kyoto-u.ac.jp

**Abstract** In peer-to-peer(P2P) content sharing, users can share their contents by contributing their own resources to one another. However, when contributing, users experience psychological dissatisfaction, namely cost. Because of the cost, most users attempt to obtain content without any contributions. To motivate users to contribute their resources to the service, incentive-rewarding mechanisms have been proposed. EMOTIVER is an incentive mechanism that accommodates users' subjective cost, which was not dealt with in conventional mechanisms. Previously, we have clarified effective rewarding methods for different dominant costs. In this paper, we assume a more diversified cost condition and evaluate the robustness of EMOTIVER in the condition. From the comparison with the ideal result, we evaluate the optimality of our rewarding methods. Furthermore, we propose a reward reassignment method and show the improvement by the method through simulation.

**Key words** P2P Content Sharing, Incentive Rewarding, Cost

### 1. はじめに

P2P 情報共有では、コンテンツ提供など参加ユーザの貢献が不可欠である。しかし、貢献行動はリソースの拠出を伴うため、心理的な負担であるコストが発生する。このため、貢献行動を行わないフリーライダーと呼ばれるユーザが多くを占めることが問題となっている [1]。そこで、ユーザの貢献に適切な報酬を与え、心理的な満足感である効用を発生させることで貢献

を促進するインセンティブメカニズムが考案されており、貢献行動に伴う主観的なコストに対応可能なインセンティブメカニズムとして EMOTIVER が提案されている [2], [3]。また、アップロードに伴う帯域の提供といった貢献実績に直結するコストと、サービスに接続しアップロードのために待機しておくことなど貢献実績に直結しない間接的なコストにコストを分類したときに、支配的なコストと有効な貢献評価関数の関係が明らかにされている [2]。しかし、近年の無線アクセス系の高速化、モ

パイル端末の高性能化を鑑みると、今後、P2P 情報共有に参加するユーザの端末形態、アクセス形態は一層多様化し、コストも同様に多様化することが考えられる。

本稿では、ユーザのコストがより多様な環境で、直接的・間接的コストの支配性がユーザの貢献度と与える影響についてシミュレーション評価を行い、コストに対する EMOTIVER のロバスト性を確認する。また、貢献評価関数の最適性を評価するため、理想的な報酬配分が行われたときの結果とも比較を行う。そして、理想値との差を補償する新たな報酬配分方法を考案する。

## 2. コストの多様性・支配性に対する EMOTIVER のロバスト性

EMOTIVER のメカニズムに関する詳細な説明は [3] に譲る。インセンティブメカニズムにおいてユーザが貢献行動に満足する条件は、報酬によりコストを上回る効用を得て、その差である利得が正となることである。このため、報酬が有限であるとき、多くのユーザの貢献を得るためには、個々のユーザのコストを精密に把握することが求められるが、これを外部から定量化するのは非常に難しい。また、同一のユーザでも環境や状況によりコストは変動する。このため、主観的なコストに対応できるよう設計された EMOTIVER でさえも、サービスに参加しているユーザそれぞれに最適な報酬の配分を行うことは非常に難しい。

しかし、直接的コスト  $C_d$  がコスト要因として支配的なユーザが多数を占める環境では実績  $r$  に応じた報酬配分を、間接的コスト  $C_i$  が支配的なユーザが多数を占める環境では、努力水準  $e$  に応じた報酬配分を行うことで、適切な報酬配分が可能であることが示されている [2]。つまり、EMOTIVER においては、ユーザの詳細なコストを把握できなくとも、ユーザ群の支配的なコストが分かれば、適切な報酬配分が可能になる。一方で、 $C_d$  が支配的なユーザと  $C_i$  が支配的なユーザが混在している環境でも、ユーザの貢献度は大きく減少しない。このことから、EMOTIVER のロバスト性も示されている。

## 3. シミュレーション評価

本節では、[2] での評価時よりも、さらにユーザの感じるコストの分散が大きい環境で評価を行い、EMOTIVER のロバスト性を確認する。

### 3.1 シミュレーションモデル

シミュレーションは、200 人のユーザがコンテンツ取得に対価を支払う P2P 情報共有に参加している状況を想定した。簡単のため、各ユーザの持つコンテンツの価値は全て等しく、アップロード要求には必ず従うものとした。本シミュレーションでは、実績  $r$  をユーザがアップロードしたコンテンツの数と定めた。そして、努力水準  $e$  を、あるサービス時間のうち、ユーザが他ユーザからのアップロード要求を受けるため待機する時間の割合と定めた。ユーザは  $e = 0, 0.5, 1$  から自身に最適な努力水準を選択する。また、ユーザはアップロードに際して直接的コスト  $C_d$  を、ある一定の待機時間に対して間接的コスト  $C_i$

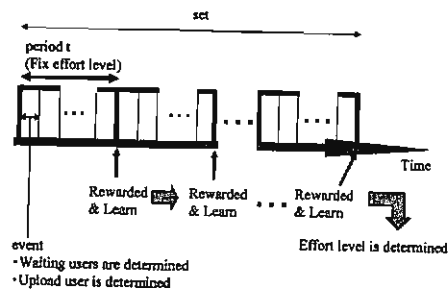


図 1 EMOTIVER シミュレーションモデル

を感じることを定めた。

EMOTIVER のシミュレーションモデルは図 1 のように表され、時間の単位は 3 種類に区別される。時間単位は小さい頃からイベント、ペリオド、セットと定めた。1 イベント毎にファイルのアップロードが行われ、1 ペリオド毎に報酬配分と学習が行われる。十分に長い時間が経ち、学習が収束するまでの期間をセットと定めた。

具体的なシミュレーションフローは以下のようになる。

- (1) 貢献行動：各々のユーザが  $e$  に基づきアップロード要求を受けるため接続待機するか否かを決める。待機ユーザには各々  $C_i$  が加算され、1 名がアップロード要求を受ける。該当ユーザはファイルアップロードを行い、 $C_d$  を被る。アップロード要求を受ける確率は、ユーザの持つコンテンツ価値が全て等しいと仮定しているため、待機ユーザ間で等確率である。一方、コンテンツを得たユーザからその対価として 1 がサービスに支払われ、報酬総和に加算される (1 イベント終了)
- (2) 報酬配分と学習： $e$  を固定して 1000 回イベント (= 1 ペリオド) を行う。1 ペリオドが終了すると、 $e$  と  $r$  から各ユーザの貢献評価値が決定され、これに基づき報酬総和が比例配分される。ユーザは受け取った報酬と感じたコストから対話エージェントに満足、不満足反応を返し、対話エージェントはユーザの反応から学習して、ユーザに新たな  $e$  を提示する (1 ペリオド終了)
- (3) 評価：ペリオドを繰り返し、学習が収束し、ユーザが最終的に選んだ  $e$  を観測する (1 セット終了)。今回は学習を十分に収束させるために、1 セット = 60 ペリオドと定めた

### 3.2 学習アルゴリズム

エージェントの学習には強化学習を用いた。強化学習とは与えられた問題環境下で試行錯誤を繰り返し、機械的に最適解を自ら見つける学習方法である。今回は強化学習の一種である Profit Sharing を用いた [4]。学習において各行動の選択肢 (本稿では努力水準) の価値の大きさを表す  $Q$  値は、現在選択している努力水準  $e$  の  $Q$  値を  $Q_e$  とすると、ユーザがそのペリオドの貢献行動で得た利得  $SAT(t)$  を対話エージェントの強化信号として

$$Q_e \leftarrow Q_e + SAT(t) \quad (1)$$

と更新される。

表1 コスト支配性の定義

支配性	$\overline{C_d}/\overline{C_i}$	$\overline{C_d}$ と $\overline{C_i}$ の相関	分布
$C_d$ が支配的	10	1	正規分布
$C_i$ が支配的	0.1	1	正規分布
ユーザごとに異なる	1	0	一様

次ピリオドにおける努力水準選択には、 $\epsilon$ -greedy法を用いた。これは、確率 $1-\epsilon$ でQ値の最も大きい努力水準を選択し、確率 $\epsilon$ でそれ以外の努力水準に遷移を行う行動決定方法である。 $\epsilon$ が定数であれば、Q値の大小が明白になった状態でも、確率的に他の努力水準を対話エージェントが提示してしまうおそれがある。これを防ぐため、今回は $\epsilon$ を $\epsilon = 0.3 - t/T * 0.3$ と定め、ピリオドが進むにつれ収束するようにした。Tは1セットのピリオド数、tは現在のピリオドを表す。

Profit Sharingでは一般に式(1)に従って、過去の強化信号SAT(t)が全て蓄積される。このため、SAT(t)が何らかの理由で極端な値を取った場合、その影響が十分小さくなるまで、エージェントがユーザに最適な努力水準を提示できない可能性がある。この問題を回避するため、今回のシミュレーションでは直近の5回分の強化信号SATのみを蓄積すると定めた。

以上のように、本メカニズムではQ値が最も高い努力水準を提示努力水準として選択する序数的効用の概念を利用して学習を行っている[5]。このため、Q値の差の大小は問題にならず、ユーザにとってどの努力水準が選好されるかさえ判明すれば、エージェントは最終的にユーザに適した努力水準を提示できる。しかし、学習の過程で最適な努力水準を見つけ出すために過去を学習する必要があり、学習を高速にするために、今回は利得SAT(t)の具体値を用いている。

### 3.3 ユーザ群のコスト支配性と貢献評価関数の設定

ユーザ群に支配的なコストが存在する場合、ユーザごとに支配的なコストが異なる場合を想定してシミュレーションを行った。ユーザのコストは多様なため、各ユーザの $C_d$ 、 $C_i$ は互いに相関のある確率分布に従って設定した。コストに支配性がある場合は正規分布を用い、 $C_d$ 、 $C_i$ 間の相関を1とした。相関を1とすることで、全ユーザで同じコストが支配的になる。ユーザごとに支配的なコストが異なる場合の確率分布は、[2]では正規分布が用いられていたが、今回は正規分布より分散が大きい一様分布を採用し、 $C_d$ 、 $C_i$ 間の相関を0とした。相関を0とすることで、支配的なコストがユーザごとに最も大きく異なる環境となる。また、いずれの確率分布を用いる場合も、[2]で行ったシミュレーションよりも大きな分散を用いた。ここで、全ユーザが $e=1$ を取るとき、各ユーザは1ピリオドで1000回(=1000イベント)貢献のための接続待機を行う。一方、各ユーザが1イベント中にアップロード要求を受け付ける確率は常に $\frac{1}{200}$ のため、1ユーザが1ピリオド中に行うコンテンツアップロード回数の期待値は5回である。よって、 $C_i$ と $C_d$ の平均値 $\overline{C_i}$ と $\overline{C_d}$ には

$$\overline{C_i} : \overline{C_d} = 1000C_i : 5C_d$$

の関係がある。上式から、表1のようにコストの支配性を定

義した。さらに、 $\overline{C_d} : \overline{C_i}$ を一定に保ったまま、コストの値を変化させた。

### 3.4 評価項目

各ユーザがどのように貢献するのが望ましいか、その評価尺度は運用ポリシーに依存する。今回は、以下の2点の評価尺度を用いた。

**努力水準期待値** ユーザが選択したeの平均値

**貢献ユーザ率** 全ユーザのうち $e > 0$ を選択したユーザの割合

努力水準期待値はユーザから全体としてどの程度の貢献が得られるかを表す評価指数である。これが高いということは、参加ユーザ1人当たりの貢献意思が高いため、アップロード要求を受け付けるため待機している時間が長いことになるので、コンテンツの入手が容易になることが期待できる。貢献ユーザ率が高くなると、フリーライダーが減少するため、アップロードされるコンテンツの種類が多様化や、アップロード負荷の分散が期待できる。

### 3.5 理想値の導出

理想的な報酬配分が実現したとき、努力水準期待値、貢献ユーザ率がそれぞれ取りうる最大値を実現可能な理想値とし、各貢献評価関数による報酬配分結果と理想値を比較する。なお、理想的な報酬配分とは、配分可能な報酬の総和が全ユーザの感じるコスト総和を上回っている条件下、 $e > 0$ を取る全ユーザの利得が正となるように、ユーザのコストに応じて報酬を配分することを指す。

理想値導出の流れは以下ようになる。

(1) ユーザの取る努力水準の組み合わせを固定し、3.1節と同様のイベントを1ピリオド分だけ行う

(2) 終了後、ユーザの利得SATの総和が、ユーザに配分される報酬をWとして、 $\sum SAT = \sum W - \sum (C_d + C_i) > 0$ を満たしているかどうかを調べる。上式が成立する場合、理想的な報酬配分を行えば、 $e > 0$ を取る全ユーザの利得をSAT > 0とすることができる

(3) 現在の努力水準期待値、貢献ユーザ率を導出し、現在までの最大値と比較して、大きい場合は更新する

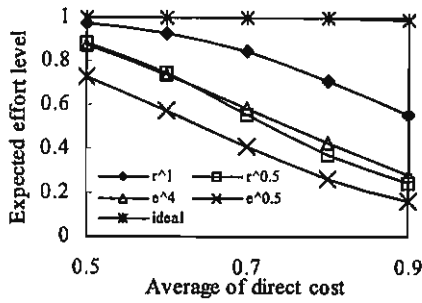
(4) ユーザの取る努力水準の組み合わせを変更し、同じ操作を取りうる全ての努力水準の組み合わせで試行する

## 4. 結果と考察

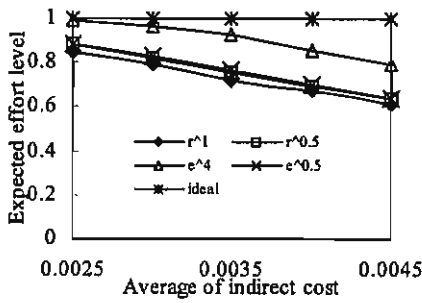
### 4.1 各貢献評価関数の比較

各貢献評価関数を用いた報酬配分を1セット行った場合の、努力水準期待値、貢献ユーザ率の結果を図2、図3に示す。図2の縦軸はeの期待値、図3の縦軸は努力水準が $e > 0$ を取るユーザの割合を示す。各図(a)、(c)の横軸はそれぞれ $C_d$ の平均値を、各図(b)の横軸は $C_i$ の平均値をそれぞれ示す。貢献評価関数の $\alpha$ の値は様々な値を取って試行したが、図には比較的有效性の高かった結果のみを示した。

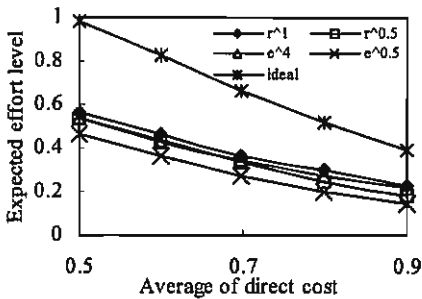
コストに支配性が存在する場合  $C_d$ が支配的なとき、図2の(a)より、 $r^1$ を貢献評価関数に用いることで、最も高い努力水準期待値が得られた。また、図3(a)のとおり、 $r^1$ またはは



(a)  $C_d$  が支配的

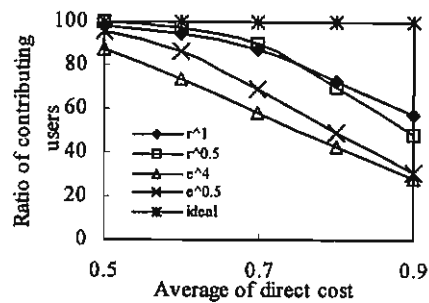


(b)  $C_i$  が支配的

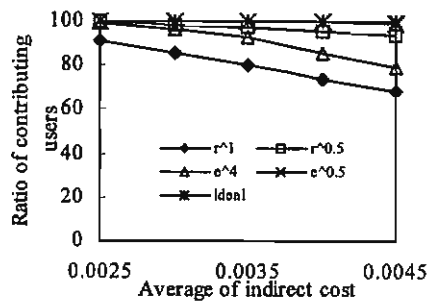


(c) 支配性多様

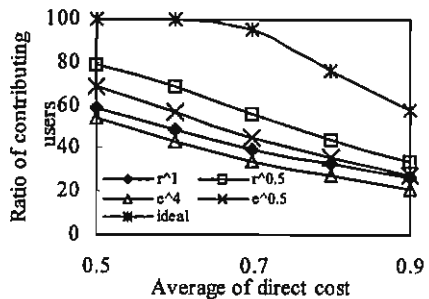
図2 努力水準期待値と理想値



(a)  $C_d$  が支配的



(b)  $C_i$  が支配的



(c) 支配性多様

図3 貢献ユーザ率と理想値

$r^{0.5}$  を貢献評価関数に用いると貢献ユーザ率が最大になった。 $C_i$  が支配的なとき、図 2(b) のように努力水準期待値は  $e^4$  を、図 3(b) のとおり、貢献ユーザ率は  $e^{0.5}$  を貢献評価関数に用いると最大となった。

以上から、 $C_d$  が支配的な場合、実績  $r$  基準の貢献評価関数が、 $C_i$  が支配的な場合、努力水準  $e$  基準の貢献評価関数が比較的有効であることが確かめられ、コストに支配性が存在するとき、支配的なコストを補償する貢献評価関数を用いることで、比較的適切に報酬配分が可能であると確認された。

ユーザごとに支配的なコストが異なる場合 図 2(c) の努力水準期待値を見ると、どの貢献評価関数を用いても、ほぼ結果が同じことが分かる。これは、 $C_d$  が支配的なユーザと  $C_i$  が支配的なユーザが混在しているため、実績基準、努力水準基準

のいずれの貢献評価関数でも、 $C_d$  ないしは  $C_i$  が支配的な一部のユーザのみを補償することになったからだと考えられる。一方で、図 3(c) のように、貢献ユーザ率は貢献評価関数の違いによって値に明確な違いが見られる。例えば、 $r^{0.5}$  と  $e^4$  を比較すると、前者では  $e=1$  のユーザ数は少ないが、 $e=0.5$  のユーザが多く、貢献ユーザ率が高い。後者では、貢献を行っているユーザの数は少ないが、ほぼ全てのユーザが  $e=1$  を選択していた。このように、ユーザごとに支配的なコストが異なるときは、貢献評価関数を使い分けることによって、ほぼ同じ努力水準期待値を保ったまま、サービスの運営ポリシーに応じて貢献ユーザの内訳を変化させることが可能である。

以上より、ユーザが感じるコストの分散を大きくしても、[2] と同様の結果が導かれたことから、コストの多様性に対する

EMOTIVER のロバスト性が確認できた。

#### 4.2 理想値との比較

コストに支配性が存在する場合 図 2, 図 3 の (a)(b) から、コストが小さいとき、最も性能の高い貢献評価関数では、努力水準期待値、貢献ユーザ率がいずれも理想値の 9 割以上の値を示していることがわかる。この理由は第一に、本シミュレーションでは個々のユーザが強いコスト支配性を持っていたため、支配的なコストを補償する貢献評価関数が非常に有効だったからであると考えられる。第二に、報酬総和に対してコスト総和が小さく、適切な報酬配分が行われていなくても、ユーザが十分に満足する報酬が配分されたからであると考えられる。しかし、コスト総和が大きくなるにつれて、それを補償するためには正確な報酬配分が必要となるため、理想値と各貢献評価関数による結果とが乖離してしまった。

ユーザごとに支配的なコストが異なる場合 図 2, 図 3 の (c) から、努力水準期待値、貢献ユーザ率ともに一貫して理想値と大きく乖離していることがわかる。ユーザごとに支配的なコストが異なる場合は、今回用いた貢献評価関数では、先述のようにそれぞれ一部のユーザしかコストを補償できないのに対し、理想的な報酬配分を行った場合、支配的なコストが異なるユーザでも補償が可能なのである。

#### 4.3 理想値に近づける方法

本シミュレーションでは、報酬は実績  $r$  あるいは努力水準  $e$  を評価する貢献評価関数に基づいて配分される。この関数は貢献評価値が等しいと、コストが大きいユーザとコストが小さいユーザに同じ報酬を配分する。この結果、コストが小さいユーザにとっては報酬が過剰、逆にコストが大きいユーザには報酬が不足という事態が起こりえる。これは特にユーザごとに支配的なコストが異なる場合に顕著であると考えられる。配分できる報酬の総和が有限であることを考慮すると、多くのユーザを貢献させる目的を達するには現在の画一的な貢献評価関数では報酬を有効に配分できないということを意味し、フリーライダーを減少させる、また、貢献ユーザの貢献意思をさらに促進するためには、より有効な報酬配分手法を検討する必要がある。

### 5. より有効な報酬配分方法の検討

前節の考察で述べたように、画一的な貢献評価関数を用いると、報酬が過剰なユーザが存在する一方、報酬が不足するユーザが存在するために、報酬配分結果が理想値から乖離することが示された。本節ではこの結果を改善するため、報酬が過剰なユーザを判別して報酬の過剰分を徴収し、コストの大きいユーザに再配分することで、貢献ユーザの努力水準を保ったまま、非貢献ユーザを減少させる報酬配分メカニズムを考案し、シミュレーションによる評価を行う。

ここでは、3.1 節で述べたシミュレーションモデルをノーマルフェーズと呼び、ノーマルフェーズ終了後に、報酬再配分により貢献ユーザの更なる増加を目指す再配分フェーズを行うことにする。すなわち、4. 節で述べた結果に対し、報酬再配分を用いることで、この結果をより改善することを考える。特に本稿では貢献ユーザ率の改善に主眼をおくことにする。

#### 5.1 徴収率を導入した報酬配分

過度に配分された報酬を、報酬が不足しているユーザに再配分するため、貢献評価関数に徴収率  $\beta (0 \leq \beta < 1)$  を導入する。徴収率を導入した貢献評価関数は、実績基準の貢献評価関数で  $(1 - \beta)r^\alpha$ 、努力水準基準の貢献評価関数で  $(1 - \beta)e^\alpha$  となる。そして、再配分フェーズでは、ノーマルフェーズで報酬が不足したために  $e = 0$  を選択した非貢献ユーザの  $\beta$  を 0 に保ったまま、努力水準  $e > 0$  を選択した報酬が過剰なユーザの  $\beta$  を大きくする。これにより、非貢献ユーザが次に貢献を行ったとき、同じ貢献に対する評価値が徴収前と比較して相対的に増加しており、配分される報酬が増加する。

##### 5.1.1 徴収率制御法

貢献ユーザの徴収率  $\beta$  を変化させる際には、徴収により貢献ユーザの努力水準が低下してしまわない範囲で、そのユーザの徴収率を最大限大きくする。

3.2 節で述べたように、各努力水準の Q 値は直近 5 回分の強化信号 SAT の和になっている。今、あるユーザにとって、最大の Q 値を持つ努力水準を  $e_{max}$ 、2 番目に大きい Q 値を持つ努力水準を  $e_{max2}$  と表すとする。そして  $e_{max}$  の Q 値を構成する 5 つの強化信号のうち、最大のものを  $SAT_{max}$  とおく。

再配分フェーズでは、ユーザが  $e_{max2}$  を選択した時にそのユーザの徴収率  $\beta$  を変化させる。そしてそのユーザが努力水準  $e_{max2}$  で貢献し、強化信号 SAT を得たとする。その際、SAT と  $SAT_{max}$  を比較し、以下のように徴収率  $\beta$  を更新する。

$$\beta \leftarrow \beta + \Delta\beta \quad (SAT < SAT_{max})$$

$$\beta \leftarrow \beta - \Delta\beta \quad (SAT \geq SAT_{max})$$

上記の式に従い  $\beta$  を更新することで、努力水準の順位を変動させない範囲で、 $\beta$  を大きくすることができる。

#### 5.2 シミュレーションモデル

再配分の有効性を確認するためシミュレーション評価を行った。再配分を行うことで、確実に非貢献ユーザを貢献ユーザへと変化させるため、徴収によって得られた報酬の余剰分を、まず、非貢献ユーザのうちのひとりに配分する。そしてそのユーザが報酬に満足し貢献ユーザになれば、次にまた非貢献ユーザをひとり選択し、そのユーザに報酬の余剰分を再配分する、という流れを採る。再配分を導入したシミュレーションの流れは以下ようになる。

(1) ノーマルフェーズ 4. 節で述べた評価と同様のシミュレーションを行う。このとき、貢献ユーザとなったユーザが徴収対象となる。

##### (2) 再配分フェーズ

1.  $\epsilon$ -greedy 法で用いる  $\epsilon$  を貢献ユーザは 0.3、非貢献ユーザは 1 に固定する。そして、ノーマルフェーズの時点で  $e = 0$  を選択したユーザから 1 名を選択し、このユーザを  $user_0$  とする。 $user_0$  には  $\epsilon$  の値に関わらず、確率 0.7 で  $e = 0.5$  を、確率 0.3 で  $e = 0$  または  $e = 1$  を提示する。これは、貢献を積極的に行うことで、貢献に満足できるかどうかを試す機会を増やすためである。その間、 $e > 0$  のユーザから徴収を開始し、上述の徴収率制御法に基づいて  $\beta$  の値を  $\Delta\beta$  ずつ増減させる。シミュレーションでは  $\Delta\beta = 0.01$  とした。

2.  $user_0$  はピリオド終了ごとに  $Q$  値を確認し、他ユーザの徴収により十分な報酬が配分されたことで、 $e = 0.5$  が最適な努力水準、すなわち  $Q_{0.5} > Q_0$  となれば  $user_0$  に対する再配分フェーズを終了する。十分な期間学習をしたにもかかわらず  $Q_{0.5} > Q_0$  とならなかった場合は、 $user_0$  に対して報酬の再配分を行う直前まで全員の徴収率をリセットし、 $user_0$  の努力水準を  $e = 0$  に戻す。

3. 引き続き別の  $e = 0$  ユーザについて同様の操作を行い、 $e = 0$  を選択した全ユーザに対して再配分が終了した後、徴収率が決定する前にユーザの取った満足・不満足の影響を排除するため、徴収率が決定した状態でノーマルフェーズを 1 セット行う。

### 5.3 結果と考察

シミュレーションの結果、 $Cd$  が支配的かつ  $\overline{Cd} = 0.8$  とコストが大きいとき、貢献評価関数として  $r^1$  を用いると、徴収対象ユーザの努力水準を減少させないまま、 $e = 0$  のユーザを  $e = 0.5$  に遷移できることが確認できた。これにより、再配分を行うことで通常の貢献評価関数を用いるよりもユーザの貢献を促進できることが確かめられた。

## 6. まとめと今後の課題

本稿では、外部評価機構と対話型学習エージェントを用いたインセンティブメカニズムである EMOTIVER のコストに対するロバスト性を示した。一方で、画一的な報酬配分結果と理想値との乖離を示した。そこで、貢献ユーザ数を増加させるため、報酬を再配分する手法の検討を行った。しかし一方で、この再配分アルゴリズムは  $Q$  値の具体的な値をエージェントが把握できると仮定して学習と徴収率の制御を行っている。また、徴収率決定のための所要時間については評価を行っていない。今後、 $Q$  値の序列のみが扱えるより現実的な環境でのシミュレーションと、再配分完了までの時間短縮方法の検討が必要である。

## 謝 辞

この研究の一部は、(財)近畿移動無線センター(モバイルワイヤレス研究助成)ならびに(財)大川情報通信基金(no.07-07)の援助を受けて実施された。

## 文 献

- [1] C.Buragohain, D.Agrawal, and S.Suri, "A Game Theoretic Framework for Incentives in P2P Systems," P2P 2003, pp.48-56, Sept. 2003.
- [2] 橋本遼, 佐藤健一朗, 新熊亮一, 高橋達郎, "無線ユーザが混在する P2P 情報共有のためのインセンティブメカニズム," ワイヤレステクノロジーパーク, May 2008.
- [3] 佐藤健一朗, 橋本遼, 新熊亮一, 高橋達郎, "ユーザのコスト多様性に対応したインセンティブメカニズム," 信学技報, MoMuC, July 2008.
- [4] 高玉圭樹, "マルチエージェント学習-相互作用の謎に迫る-, " コロナ社, 2003.
- [5] Hal R.Varian, "入門ミクロ経済学," 勁草書房, 2007.