ポリシーランキングに基づくプライバシポリシー交渉方式

古川　諒†　　川戸　正裕†　　伊東　直子†　　中江　政行†

†NEC サービスプラットフォーム研究所
211-8666 神奈川県川崎市中原区下沼部 1753

{r-furukawa@cb, m-kawato@ap, naoko@cj, m-nakae@bp}.jp.nec.com

あらまし　近年，種々のプライバシ情報を収集しユーザに適した情報をフィードバックするサービスが注目されている．このようなサービスにおいては，ユーザのプライバシ情報保護のために，ユーザによる情報の制御が必要である．このために，P3P や Prime などのプロジェクトにおけるポリシーを用いた制御技術や，ユーザ・サービスプロバイダ双方の要件を満たすためにポリシー交渉技術が開発されてきた．しかしポリシーが競合する時には，ユーザのポリシーとはかけ離れたサービスプロバイダのポリシーに合わせるか，あるいは多数回の交渉が必要でコストが高いという問題があった．本稿ではこの問題を解決するためにポリシーランキングに基づくポリシー交渉方式を提案する．そして，対話回数とポリシーの柔軟性，及びポリシー交渉に必要な時間，ポリシーランキングの妥当性について評価を行い提案方式の有効性を示す．

A Privacy Policy Negotiation Method Based on Policy Ranking

Ryo Furukawa†　　Masahiro Kawato†　　Naoko Ito†　　Masayuki Nakae†

†Service Platforms Res. Labs., NEC Corporation
1753, Shimonumabe, Nakahara-Ku, Kawasaki-Shi, Kanagawa, 211-8666, Japan,

{r-furukawa@cb, m-kawato@ap, naoko@cj, m-nakae@bp}.jp.nec.com

Abstract　Recently, more and more services collect many user's privacy information and feedback personalized contents to users. With these services, it is essential that users can control how their privacy should be handled to protect user's privacy information and several policy based approaches have been developed. However, in the policy based approach, user's privacy policy often conflicts with service provider's privacy policy. To solve this problem, policy negotiation methods based on the dialogue between user and service provider have been proposed. But some existing methods only allow choosing a policy set by a service provider and users may need to largely compromise their policy. Other methods require a number of dialogues to come to an agreement. In this paper, we propose and evaluate the policy negotiation method based on policy ranking, which allows agreeing a policy from a wide variety of policies with a minimum number of dialogues for the negotiation.

1　はじめに

近年，NTT ドコモが提供する i コンシェルや Amazon における商品推薦サービスなど，ユーザの位置情報や購買情報といったプライバシ情報を収集し，活用するサービスが注目されている．これらのサービスにおいて，収集される情報はプライバシ情報であり，ユーザが情報の流れや取り扱いなどを適切に管理できる必要がある．

このようなプライバシ情報管理をポリシーベースで行う技術として，P3P[1] や Prime[2] といったプロジェクトにおいてプライバシポリシーに関する研究が取り組まれてきた．

これらの研究において，プライバシポリシーはユーザのプライバシ情報提供の可否の指定や，情報提供に関する条件付けなどといった保護要件を記述したり，サービスプロバイダがプライバシ情報を収集に際して，どのような情報の取り扱いを行うかといった利用要件が記載される．プライバシポリシーを用い，ポリシーが遵守されることによって，ユーザは安心してプライバシ情報を提供し，サービスを受けることができる．

しかし，ユーザの保護要件とサービスプロバイダの利用要件の間には，２つの要件を同時に満たせない競合が発生する可能性がある．たとえば，ユーザは位置情報を提供したくないが，サービスは位置情報を取得しないとサービスを提供できないため，必ず取得する必要がある時などには，ユーザのポリシー

とサービスのポリシーの間で競合が発生してしまう．ポリシー間の競合を解消しなければプライバシ情報の取り扱いを決定できず，サービスを受けることはできない．ポリシーを妥協することでサービスを受けられるようにすることが望まれる場合には，競合を効率的に解決する必要がある．

ポリシー間の競合を解決する方法として，ポリシー交渉技術が文献 [1][3] などで提案されている．ポリシー交渉技術は，ユーザとサービスプロバイダ間の対話によって，競合を解決する技術である．しかし，これらの既存手法は，ユーザがサービス側で一方的に定めた唯一のポリシーに同意しなければならず，ポリシーがフレキシブルに設定できなかったり，フレキシブルに設定するには，ユーザやサービスプロバイダが何度も対話を繰り返してポリシーを細かく調節していく必要があるため非効率的であるなどの問題があった．

そこで本稿では，ユーザとサービスプロバイダ間で一度の対話しか行わずに，フレキシブルなポリシーを設定可能なポリシー交渉方式を提案する．また，提案方式を用いた場合にポリシー交渉にかかる時間を計測し，実サービスとして問題ない速度でポリシー交渉が可能であることを示す．そして，提案方式によって一度の対話だけでも妥当性の高いポリシーを設定可能であることを示す．

2　ポリシー交渉技術と課題

2.1　プライバシポリシー

プライバシポリシーは，プライバシ情報の取り扱い方や提供の可否などの要件を定めるルールの集合である．P3P や Prime 等で定義されたプライバシーポリシーでは，一つのポリシーに氏名・年齢等といったプライバシ情報の属性と，それに対しての提供の可否や提供条件といった取り扱いの要件であるアクションが記載される．

本稿では属性を dt，アクションを a とし，この 2 つの組 (dt, a) が一つのルールとする．ルールの集合であるプライバシポリシー Pol は以下のように記される．

$$Pol = \{(dt, a)\}$$

2.2　プライバシポリシー交渉技術

実際にサービスプロバイダがプライバシ情報を扱う場合，ポリシーはユーザが求める保護要件（ユーザポリシー Pol^u）と，サービスプロバイダが求める利用要件（サービスポリシー Pol^s）の 2 つのプライバシポリシーを参照してプライバシ情報の取り扱いを決定する必要がある．

しかし，この 2 つのポリシーに設定されたルールの間にはしばしば競合が発生する．ポリシー間で競合がある場合，ユーザとサービスプロバイダ間でプライバシ情報の取り扱いを決めることができず，サービスを提供できなくなってしまう．このため，ポリシー競合を解消する，ポリシー交渉技術 [1][3] の研究が近年行われている．

ポリシー交渉技術とは，図 1 に示すように，ユーザとサービスプロバイダの間で交渉を行い，どちらかのポリシーを変更しながらユーザポリシーとサービスポリシー間の競合を解消することで，両者が合意するポリシー（合意済みポリシー）を生成する技術である．

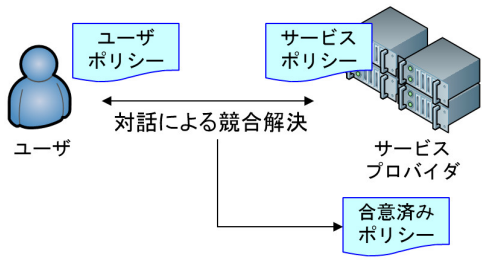

図 1: ポリシー交渉

2.3　既存技術と課題

ポリシー交渉技術には代表的な 2 つの手法が提案されている．

第一の手法は P3P[1] において利用される方法である．この手法は，ユーザポリシー，サービスポリシー間の競合を判定し，競合が発生した場合にのみ，ユーザ側にサービスポリシーに対する許諾を求めるというものである．ユーザの許諾が取れた場合にサービスポリシーを合意済みポリシーとし，取れなかった場合にはユーザがサービスを受けられないようになる．本手法は対話は 1 度で終わるため交渉コストは低いが，ユーザは唯一のサービスポリシーに同意することしかできないため，フレキシブルにポリシーを設定できないという欠点がある．

第二の手法はユーザとサービスプロバイダ間で，対話をしながらポリシーの修正を行い，競合が無くなるまで対話を繰り返す手法である [3]．本手法は，前者と異なりフレキシブルなポリシー設定が可能であるが，ユーザやサービスプロバイダ間で何度も対話が必要である．

交渉における対話回数の増加は，ユーザにとっては，ポリシー交渉に必要な時間の増大，サービスプロバイダーにとっては，個々のユーザに対して対話のたびにポリシーを修正する必要性が発生することになり，交渉コストが高くなってしまう．

このように，既存のポリシー交渉方式は，ユーザとサービスプロバイダ間の対話回数の低減と，生成される合意済みポリシーの柔軟性を両立できない．そこで，本稿では，最低限の対話で柔軟なポリシーを生成できるポリシー交渉方式の設計を行う．

3 ポリシーランキングを用いたポリシー交渉方式

本節では，前節で述べた課題を解決する，ポリシーランキングを用いたポリシー交渉方式を提案する．

まず，課題を解決するために，以下のような方式の設計指針を立てた．

方式設計指針

- 柔軟なポリシーを低コストで設定するために，複数のポリシー修正候補が事前に準備されるべきである
- 対話回数低減のために，一度の対話で複数のポリシー修正候補が示されるべきである
- 複数のポリシー修正候補の中で，ユーザに適したポリシーが明確であるべきである

□

以上の設計指針を満たすために，ポリシーランキングを用いたポリシー交渉方式を提案する．

本方式は，ポリシー修正候補として，複数のサービスポリシーを登録し，サービスポリシーの中からの選択基準として，ユーザポリシーに対する乖離度を基準としたポリシーランキングをユーザへ提示する．ユーザは提示されたランキングの中から，自分に適したポリシーを選択することとなる．

本節ではまず，ポリシーランキングの生成方法について 3.1 項で述べ，その中で用いるポリシー乖離度の計算方法を 3.2 項で示した後，それを用いたポリシー交渉方式について 3.3 項で述べる．

3.1 ポリシーランキング生成方法

前述したように，ポリシーランキングはユーザポリシーに対する乖離度を基準としたサービスポリシーのランキングである．

ポリシーランキングは，ユーザにとって望ましいポリシーが上位にランクされるべきだと考えられる．ユーザの要件はユーザポリシーであるため，これにより近いポリシーがより望ましい．ここで，ポリシーの近さとは，属性ごとにアクションに記載される保護要件と利用要件の近さに基づく．

また，ポリシー交渉はポリシー間の競合を解決する方法であるが，提案方式においては事前に設定されたサービスポリシーの中に，ユーザポリシーと競合が発生しない場合がある．このような場合でも，最終的なポリシーの選択はユーザが行うことが望ましいため競合がある場合と同様に扱い，ランキングとして出力する必要がある．

以上の点から，ユーザにとって最適なサービスポリシーを選択可能なランキングを出力するために，以下のランキング設計指針を立てた．

ランキング設計指針

- ユーザポリシーと競合のないポリシーが上位にランクされる
- ユーザポリシーに近いポリシーが上位にランクされる

□

この指針を満たすため，ポリシー間の近さを比較でき，かつ競合の有無が判定できる評価指標として，ポリシー乖離度を導入する．ポリシー乖離度は以下のように定義する．

定義 1：ポリシー乖離度

ユーザポリシー Pol^u とサービスポリシー Pol^s 間のポリシー乖離度 $dis(Pol^u, Pol^s)$ は，以下の性質を満たす．

$dis(Pol^u, Pol_1^s) < dis(Pol^u, Pol_2^s)$ ならば，Pol_1^s は Pol_2^s と比較して Pol^u に近い

$$dis(Pol^u, Pol^s) > 0,\ if\ \ Col(Pol^u, Pol^s) = 1$$

$$dis(Pol^u, Pol^s) < 0,\ if\ \ Col(Pol^u, Pol^s) = 0$$

□

ここで，$Col(Pol^u, Pol^s)$ は 2 つのポリシー間で競合が発生する場合に 1 を返し，発生しない場合に 0 を返す関数である．

以上のような性質を持つポリシー乖離度を用いて，ランキング設計指針を満たすポリシーランキング生成のアルゴリズムを以下に示す．具体的なポリシー乖離度計算方法は次節にて述べる．

ポリシーランキング生成アルゴリズム

1. 全てのサービスポリシー Pol_i^s に対して，ユーザポリシー Pol^u に対するポリシー乖離度 $dis(Pol^u, Pol_i^s)$ を計算する
2. $dis(Pol^u, Pol_i^s) \leq 0$ を満たす Pol_i^s を $|dis(Pol^u, Pol_i^s)|$ が小さい順にソートし，$PR_- = (Pol_1^{s-}, ...Pol_k^{s-})$ とする
3. $dis(Pol^u, Pol_i^s) > 0$ を満たす Pol_i^s を $|dis(Pol^u, Pol_i^s)|$ が小さい順にソートし，$PR_+ = (Pol_1^{s+}, ...Pol_l^{s+})$ とする
4. 2 つのランキング PR_- と PR_+ を結合してポリシーランキング $PR = (Pol_1^{s-}, ..., Pol_k^{s-}, Pol_1^{s+}, ..., Pol_l^{s+})$ として出力する

このポリシーランキング生成方法は，競合がないポリシーの中でユーザポリシーに近いものが優先さ

れ，競合があるポリシーの中でもユーザポリシーに近いものが優先されるようにランキングが構成される．そのため本方式は，ユーザにとって重視すべきサービスポリシーをランキング上位にランクすることができる．

3.2 ポリシー乖離度計算方法

本稿で利用するポリシー乖離度計算方法について述べる．

まず，ポリシー乖離度を数値的に求めるために，ポリシーを数値的に比較可能な形へ変換することを考える．

その方法として，ポリシー $Pol = \{(dt, a)\}$ を各属性 dt を一つの次元としてアクション a をプライバシ保護強度を表す数値へ変換することで，実数値ベクトル $PV \in R^n$ へ写像する方法を用いる．ここで，n は属性の数である．アクションの数値変換の具体例は 4.1 項で示す．

次に，写像により得られた実数値ベクトルを用いて乖離度の計算を行う．乖離度は基本的には，ユーザポリシーとサービスポリシー間のプライバシ保護強度の差で計算するが，競合が発生する場合には，競合する次元のみについて差を取るようにしている．これは，競合するポリシーの乖離度を比較するとき，競合しない属性よりも競合する属性に着目するためである．

また，実数値ベクトル上では，ユーザポリシーとサービスポリシーの任意の次元の値 v_i^u, v_i^s が $v_i^u > v_i^s$ を満たすとき，つまりユーザの望むプライバシ保護強度のほうがサービスプロバイダが望むプライバシ保護強度より高い場合に競合が発生するとした．

以下に，ポリシー乖離度計算方法を示す．

ポリシー乖離度計算方法

1. Pol^u, Pol^s 間で競合があるかどうかを判定する
2. 競合がある場合 $dis(Pol^u, Pol^s) = \sum_i^n d(v_i^u, v_i^s)$ とする．ただし，$d(v_i^u, v_i^s) = |v_i^u - v_i^s| \quad if \ \ v_i^u > v_i^s$,$d(v_i^u, v_i^s) = 0$ $if \ \ v_i^u \geq v_i^s$ である
3. 競合がない場合，$dis(Pol^u, Pol^s) = \sum_{i=1}^n -|v_i^u - v_i^s|$ とする

このようにすることで，定義 1 を満たし，プラバシ保護強度の大小を反映したポリシー乖離度を計算することができる．

3.3 ポリシー交渉方式

ポリシーランキング生成方法を用いた，提案方式の動作概要を図 2 に述べる．

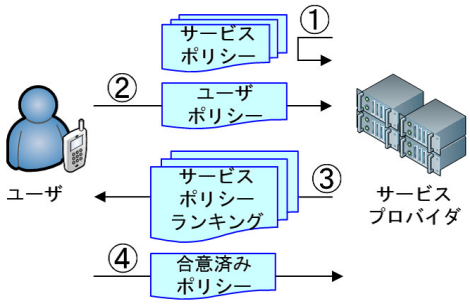

図 2: 方式の動作概要

また，本方式のアルゴリズムを以下に示す．

ポリシー交渉方式概要

1. サービスプロバイダが事前に許容できるサービスポリシーのリストを準備しておく
2. ユーザがユーザポリシーを記述し，サービスプロバイダへ送信する
3. ユーザポリシーとのポリシー乖離度を基準としたサービスポリシーランキングをユーザへ送信する
4. サービスポリシーランキングからサービスポリシーを選択し，合意済みポリシーとしてサービスプロバイダへ送信する

本方式を用いることにより，ユーザがサービスプロバイダと一度だけ対話を行えば，事前に設定されたサービスポリシーの中からユーザのポリシーを満足する，または近いポリシーを設定できる．このように，本方式は，対話負荷の低減と柔軟なポリシー設定の両立を実現できる．

4 評価

本節では提案したポリシー交渉方式について評価を行う．評価は，提案手法の対話回数とポリシーの柔軟性について，及び，ユーザとサービスプロバイダ間のインタラクションにおいて重要となる処理時間の測定と，ユーザにとってのランキングの妥当性の検証を行った．

4.1 実験設定

4.1.1 ポリシー形式

本実験で対象とするポリシーの形式は，ユーザポリシー，サービスポリシーそれぞれに，属性 dt と，それに対するアクション a の組が複数記述される．そしてアクションには，“提供（取得）しない”，“保護して提供（取得）する”，“保護しないで提供（取

得）する”のいずれかが記述される．さらに”保護して提供”する場合には，パラメータとして，プライバシ保護強度を “弱”，“中”，“強”の３段階で指定できる．プライバシ保護の強度はたとえば，$k-$ 匿名化 [4] の場合は k の値に関係したり，暗号化であれば，暗号の鍵長に関係するものである．

このポリシーにより，サービスプロバイダーは，ある属性に対して，そのデータを取得しなかったり，保護を望ましい強度で行ったり，保護しないで取得するといった意思決定を行うことができる．

4.1.2 ポリシー写像

本実験で用いたポリシーの実数値への写像方法について述べる．

前述した，アクションの値を表 1 にしたがって，実数値へ変換することで，写像を行う．

4.2 対話回数とポリシーの柔軟性の評価

本研究の目的であった，対話回数の低減とポリシーの柔軟性について評価を行う．

対話回数については，ユーザとサービスプロバイダ間で一度行うことによりポリシー交渉が可能であり，対話回数を低減できている．

表 1: アクションの実数値への写像対応表

アクション	パラメータ	変換値
保護しない	-	0
保護する	弱	1
保護する	中	1.5
保護する	強	2
提供（取得）しない	-	3

ポリシーの柔軟性については，サービスプロバイダ側で用意されるサービスポリシーのバリエーションに左右されるが，サービスプロバイダが許容できる範囲のポリシーを用意することにより，複数回対話する場合と同等の柔軟性を確保できる．

サービスプロバイダーがポリシーを複数用意する負荷が高くなるが，一人ひとりのユーザと対話を行う負荷に比べ，一度のみの負荷であり，最終的には負荷は低減されると予想している．

4.3 処理時間の評価

本項では，ユーザにかかる交渉コストの一つであるポリシー交渉にかかる処理時間の評価を行う．

ここでは，適当に生成したポリシーの数を $\{100, 200, ..., 1000\}$ と変化させてユーザのポリシー入力から，ランキングが出力されるまでの実時間を計測した．また，個々のポリシーはそれぞれデータ属性に対するルールを 10 個持つように生成した．計測結果を図 3 に示す．

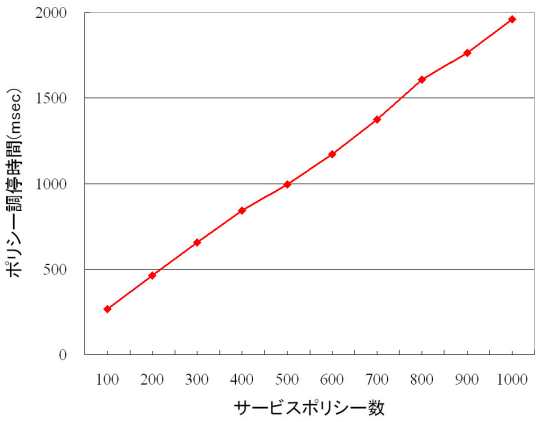

図 3: ランキング生成時間

図 3 に示すように，ポリシー数が増加するにつれ，ポリシーランキング生成処理にかかる時間が線形に増加することがわかる．

また，収集する属性の数が 4 つのサービスがあり，一つの属性に対して記述できるアクションが 5 種類あった場合，全てのパターンのポリシーを作成したとしても，$5^4 = 625$ 個であり，この程度のポリシー数を扱えれば十分であるといえる．Forrester のホワイトペーパー [5] によれば，ユーザが許容できるレスポンスタイムは 2 秒（2009 年時点）であると記載されている．提案方式では 2 秒のレスポンスタイムで 1000 以上のポリシーを処理可能であり，十分多くのポリシーをユーザが許容できる時間内に処理できているといえる．

4.4 ポリシーランキングの妥当性評価

提案方式においてユーザが自分にとって望ましいポリシーを設定可能か調査するために，提案方式により生成されるポリシーランキングの妥当性について評価する．

評価方法として，16 個のサービスポリシーを作成し，それに対する 4 つのユーザポリシーを入力とした理想的なランキング $Rank^{id}$ を構築し，提案手法による結果 $Rank^{p}$ と平均的にどれほどずれがあるかを評価した．

ランキングのずれを評価するために，検索の評価に利用される再現率と適合率に基づく指標により評価を行った．本稿で利用する評価指標は，$Rank^{id}$ の k 位以上に含まれる集合 $Rank^{id}_{k}$ を正解集合とし，$Rank^{p}$ の k 位以上に含まれる集合 $Rank^{p}_{k}$ を検索集合とした場合の適合率を計算する．なお，このように集合を定義した場合，適合率と再現率は同値となる．この適合率は以下の式で表される．

$$Prec(k) = \frac{|Rank_k^{id} \cap Rank_k^p|}{k}$$

k を変動させたグラフおよび，ランダムにランキングを作成した場合の期待値を図 4 に示す．

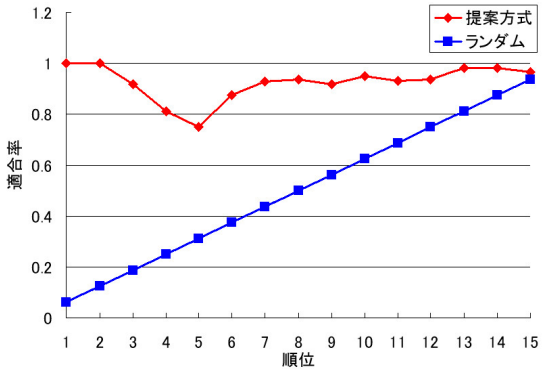

図 4: ランキングの適合率

結果から，ランダムにランキングを生成した場合と比較して，提案方式により生成されるランキングは，理想的なランキングに十分近いことがわかる．また，最悪でも適合度が約 0.78 であり，最大でも 1/5 程度しかずれが生じないことがわかる．これにより，ユーザは提示されたランキングから望ましいポリシーを選択可能である可能性が高い．

また，各ポリシーの 2 つのランキングにおける順位の近さに着目した評価を行った．この評価は，それぞれのランキングにおける，あるポリシーの順位を $r^{id}(\cdot), r^p(\cdot)$，ポリシー集合を P としたとき，以下の式で表すことができる．

$$diff = \frac{\sum_{p \in P} |r^{id}(p) - r^p(p)|}{\sum_{i=1}^{|P|} ||P| + 1 - 2 * i|}$$

この評価値を提案方式に対して計算すると，ランダムなランキングの場合の期待値 0.5 と比べ，0.12 という非常に小さいものであった．このことから各ポリシーの順位間には大きな差が無く，2 つのランキングの差異は局所的であることがわかる．図 4 の $k = 6, 7, 8$ などにおいて，直前の順位より適合率が改善する現象が起きているが，これは局所的にのみ順位が異なるため，ランク k で適合率が多少悪くても，$k+1$ まで見ると $Rank_k^{id}$ にすでに入っているポリシーが追加されるためであるといえる．

また，実際にランキングとポリシー乖離度を見ると，提案方式は同じポリシー乖離度を持つサービスポリシーが多くなる傾向にあった．そして，ランキングの局所的な差異は，この同一の乖離度を持つポリシーに対するランキングの差であることが多かった．より適切なランキングの生成にはこのような同一乖離度を持つポリシーに対して適切な順位付けを可能とする必要があるといえる．

5 おわりに

本稿では，ユーザが望むプライバシ情報に対する保護要件と，サービスプロバイダによるプライバシ情報の利用要件という 2 つのプライバシポリシーの間で発生する競合を解決する，ポリシー交渉方式を提案した．

提案方式では，許容できるサービスポリシーを複数用意し，ユーザポリシーに基づいたランキングを用いてユーザのポリシー選択を容易とすることで，一度の対話で，フレキシブルなポリシー設定を可能としている．また，提案方式においてポリシー交渉にかかる時間が十分短いことを示した．さらに，提案方式では，妥当性の高いポリシーランキングが出力されるが，同程度乖離したポリシーの順位付けに問題があることを示した．

今後の課題は提案方式が出力するポリシーランキングがユーザにとって望ましいかを，アンケート等を用いた方法により，より正確に評価してポリシーランキング生成方法を改良することである．さらに，サービスポリシーの生成支援や，ユーザポリシーを正確に記述するための支援などが必要であると考えている．

謝辞

本研究の一部は，総務省託研究「平成 22 年度大規模仮想化サーバ環境における情報セキュリティ対策技術の研究開発」の成果である。

参考文献

[1] W3C, “Platform for Privacy Preferences (P3P) Project”, `http://www.w3.org/P3P/`.

[2] “PRIME - Privacy and Identity Management for Europe”, `https://www.prime-project.eu/`.

[3] M. Hatakeyama, H. Gomi, “Privacy Policy Negotiation Framework for Attribute Exchange”, W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 2006.

[4] L Sweeney, “k-anonymity: a model for protecting privacy”, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems: 10(5), pp.555-570, 2002.

[5] Forrester, “eCommerce Web Site Performance Today”, Forrester white paper, 2009.