

QoSとプライバシーコストを考慮したWebサービス連携手法の提案

A Proposal of Service Composite Based on QoS and Privacy Cost

新井翔大*¹

Shota ARAI

菱山玲子*¹

Reiko HISHIYAMA

*¹早稲田大学創造理工学研究科経営システム工学専攻
Graduate School of Creative Science and Engineering, Waseda University

In recent years, The business model has been changed due to the popularity of the Internet spread and the web service increases. Web service has been used as a base supporting the system of the company. If web service is used widely, the web service increases and the kind of web service diversifies, a web service do not satisfy their wants. It is necessary for us to the construct the new service that some web services are cooperated. In this study, it is aimed for the construction of the service cooperation model that considered the quality as the multi-service and protection for the privacy data that the user wants to protect. As the result of simulation experiment, the services replace and new services extract for the user changing the privacy policy of the provider, and the construction of the service cooperation model is showed the effectiveness of the model.

1. 研究背景と目的

インターネットの普及やWWW(World Wide Web)の進展により、通信技術やコンピュータの処理形態が変化し、情報システムの相互連携への要求が高まっている。特にインターネットを介した企業のシステム管理や商取引は企業のビジネスモデルやシステム構築方法に大きな影響を与えている。今日では企業のシステムを支える基盤としてWebサービスは必要不可欠なものになっていると考えられている。今後Webサービスが広く利用され、サービスを提供するプロバイダが増加し、それに伴いWebサービスの増加、その種類も多様化することが予想される。この環境においてユーザは一つのサービスでは要求を満たすことが出来ず、インターネット上の様々なWebサービスを用いて複数のサービスを連携させた新しいサービスの構築が必要になってくると考えられる。また、サービスの複雑化によりサービスの質だけではなく信頼性や安全性などを考慮する必要がある。

本研究では、複数の単体サービスを連携することで生成される複合サービスとしてのサービス品質と、ユーザが守りたいプライバシー情報に対する補償を同時に考慮した、サービス連携モデルの構築を目的とする。

2. 問題の概要

2.1 Webサービスのオーケストレーション

Webサービスのオーケストレーションとは、図1のようなサービスのワークフローが与えられ、ワークフローの抽象サービスに単体サービスの候補集合から単体サービスを選ぶ手法であり、複合サービスの構築に用いられる。ここで、ワークフローは処理手順を表し、抽象サービスとは、具体的なサービスとは異なり枠組みやその概要だけが決められたサービスを表し、単体サービスは抽象サービスの特徴をもった実際にある具体的なサービスを表している。また、複合サービスは複数の単体サービスを連携させるサービスである。サービスオーケストレーションでは、サービスの候補集合からどのようにして単体サービスを選び、複合サービスの構築を行うかが重要である。

連絡先: 新井翔大, 早稲田大学大学院創造理工学研究科, s-5626.arai@akane.waseda.jp

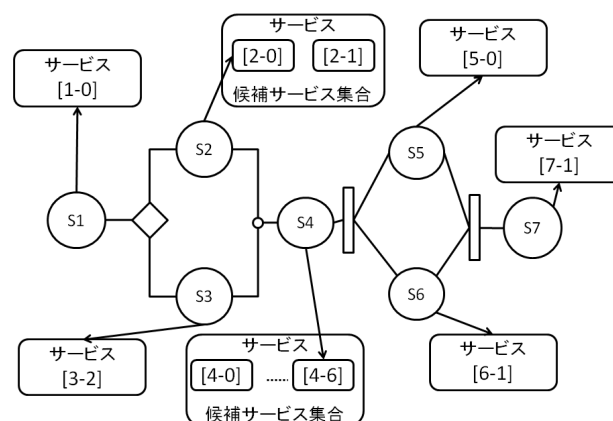


図 1: 単体サービス選択

2.2 Webサービス連携の事例と問題

Webサービス連携の事例としてオンライン上で旅行計画を行うことを考える。飛行機予約サービスやホテル予約サービスなど一つのWebサービスだけでは旅行計画を一度の手続きで行うことができないため、各々のWebサービスを連携させた複合サービスの構築が必要である [青木 2005]。複合サービス構築のためのWebサービスの選定では、Zengら [Zeng 2004] はWebサービスの品質を表すQoS(Quality of Service)のスコアリングを用いた研究を行っている。しかし、QoSだけではユーザやプロバイダの個人情報を扱い方を示すプライバシーポリシーが考慮されず、プライバシーデータが望ましくない方法で公表されるなどの危険性が高くなってしまふ。

本研究ではまず効率よく複合サービスを構築するために遺伝的アルゴリズムを用いる [Canfora 2005]。さらにこれらのプライバシー問題に対しては、QoSのみだけではなくユーザとプロバイダのプライバシーポリシーを考慮し、遺伝的アルゴリズムの評価関数に加える。ユーザのプライバシーポリシーはYuら [Yu 2006] が提案したプライバシーコストを新たな評

価指標として用いる。プロバイダのプライバシーポリシーは Hewett ら [Hewett 2009] の研究より、サービスのデータの出力に着目し、強度が異なる制約を加える。

3. 関連研究

3.1 遺伝的アルゴリズムを用いたサービス選択

Canfora ら [Canfora 2005] は遺伝的アルゴリズムを用いたサービス連携のためのサービス選択の提案した。以下にサービス選択への遺伝的アルゴリズムの適用方法とそのアルゴリズムを示す。

Step1) ランダムに N 個の個体を生成する。遺伝子の長さはワークフローの抽象サービスの数とする。遺伝子一つ一つにサービスをあてはめることによりゲノム $g=(1, \dots, N)$ を生成する。世代を T とし、 $T=1$ とする。最大世代数 T_{max} になるまで繰り返す。以下の図 2 にコーディング方法を示す。

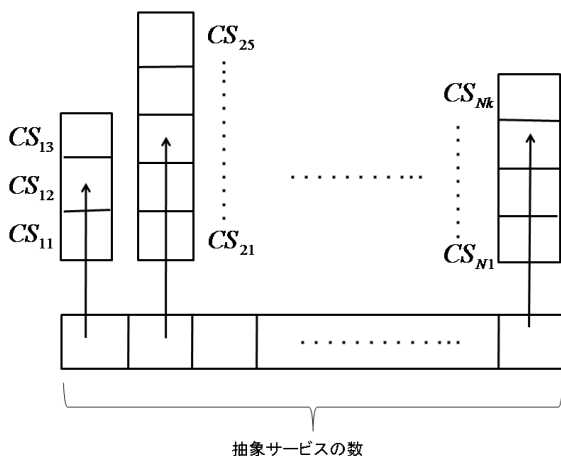


図 2: コーディング

Step2) 評価関数を用いて、現世代の各個体の適応度 $F(g)$ をそれぞれ計算する。評価関数はコスト、レスポンスタイム、信頼性、稼働率、制約で定義される。複合サービスの QoS は制約の集合 $cl_i(g)$ を持っているので、制約集合は以下のように定義する。

$$cl_i(g) \leq 0, i = 1, \dots, n. \quad (1)$$

制約充足の距離を式 (2) のように定義し、これを評価関数に用いる。

$$D(g) = \sum_{i=1}^n cl_i(g) \cdot y_i \quad (2)$$

制約が満たされているならば y_i は 0, 満たしていないならば 1 となる。式 (1), 式 (2) と QoS を用いた評価関数は以下ようになる。

$$F(g) = \frac{w_1 Cost(g) + w_2 ResponceTime(g)}{w_3 Availability(g) + w_4 Reliability(g)} + w_5 D(g) \quad (3)$$

適応度 $F(g)$ が高いほど適したサービスを選択していることになる。ここで、 w_1, \dots, w_5 は重みである。QoS の指標項目については、 $[0, 1]$ に標準化を行う。

Step3) エリート戦略によって適応度の高い遺伝子 2 個が選択される。残りは個体の適応度に応じてルーレット戦略を用いて、選択が行われる。

Step4) 交叉は 2 点交叉あるいは 1 点交叉を用いる。

Step5) 突然変異により、ある個体の遺伝子にあるサービスを他の個体の遺伝子のサービスと入れ替える。 $T=T+1$ として Step2 に戻る。

3.2 個人情報考慮したサービス選択

近年、ユーザの個人情報を考慮した Web サービス連携のためのサービス選択が行われている。Hewett ら [Hewett 2009] や Xu ら [Xu 2006] はサービスの個人情報の入出力を登録したレポジトリからサービスの入出力関係に着目し、連携可能なサービスの選択を行うモデルを提案した。しかし、サービスの品質を表す QoS が用いられておらず現実的ではない。また、Yu ら [Yu 2006] はユーザとサービスプロバイダのプライバシーポリシーに着目し、新たな評価指標であるプライバシーコストを提案している。ただし、これらの研究はプライバシーデータを企業間の共有はユーザの同意を前提とする。

3.3 プライバシーコスト

ホテルの予約サービスなどユーザの個人情報であるプライバシーデータを用いる Web サービスの利用が高くなっている。これらの Web サービスを実行するとき、プロバイダはよりユーザの要求に合うサービスを提供するためにユーザの望まない方法でプライバシーデータを収集、公表をすることがある。これは、ユーザの個人情報が危険にさらされている可能性が高いと考えられる。この背景から、ユーザのプライバシーリスクを定量化し、Yu ら [Yu 2006] はプライバシーコストを提案した。ここでの、プライバシーリスクはユーザが望まないにもかかわらず、第三者によって当核ユーザの個人情報が収集ないし公表・公開されてしまう危険性を示す。

4. 提案

4.1 提案の背景

Web サービスの増加、多様化により、ユーザに適したサービスの構築が必要になっている。そのサービスを構築するために、複数のサービスを連携することが重要であると考えている。また、Web サービスではプライバシー保護のためにユーザとプロバイダともにプライバシーポリシーを持っており、複合サービスにおいてもそれらを反映する必要がある。そこで、本研究では従来から用いられていた QoS に加えて、ユーザとプロバイダのそれぞれのプライバシーポリシーをサービス連携の際に取り入れたモデルを構築し評価する。

4.2 提案概要

本研究では、ユーザとプロバイダのプライバシーポリシーを考慮し、ユーザに適した Web サービス連携を行う。評価関数にプライバシーコストと QoS を用いて、さらにサービス間のプライバシーデータの入出力に強度の異なる制約を加える。この新たな評価関数を用いた遺伝的アルゴリズムにより、ユーザに適したサービス連携が可能になり、これらを満たすブローカエージェントモデルを提案する。ただし、プライバシーデータの企業間の共有はユーザの同意を前提とする。

4.3 提案モデル

本研究で提案するブローカーエージェントのモデルの全体図を図3に示し、そのエージェントのアルゴリズムを以下に示す。

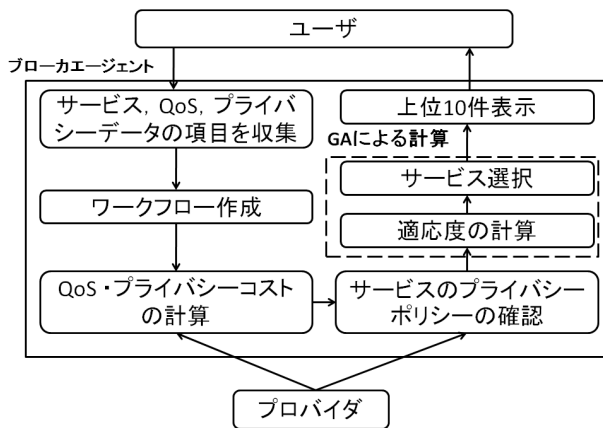


図 3: 提案モデルの全体図

- Step1) ユーザがサービス、QoS・プライバシーデータの項目を決める。
- Step2) QoS・評価項目をもとにワークフローを決める。
- Step3) QoSの標準化、プライバシーコストの計算、サービスのプライバシーポリシーの確認をする。
- Step4) 遺伝的アルゴリズムを用いて適応度計算し、サービス選択を行う。
- Step5) 比較的最適な解の上位10件を表示する。

4.4 プライバシーコストの算出

電話番号やEmailなどのプライバシーデータはまずそれぞれのカテゴリに分けられ、カテゴリ毎にユーザが公表してもよいプライバシーデータ量とサービスに含まれるデータ量の割合を計算する。それらにカテゴリのウェイトを考慮し、それらの総和がプライバシーコストとなる。以下にプライバシーコストの計算方法を示す。

$$R_i = \frac{\alpha_i}{Z_i} \cdot W_{ci} \quad (4)$$

R_i はカテゴリ i に対するプライバシーリスクである。 α はサービスに含まれるプライバシーデータ量、 Z_i はリクエストの公表してもよいプライバシーデータ量である。

$$F_p = \sum_{i=1}^k R_i \quad (5)$$

F_p はプライバシーコストでありプライバシーリスクの総和である。 k はカテゴリ数を表し、本研究では $k = 3$ である。

Yuら [Yu 2006] では、サービスの入力のみ着目していたが、本研究では、サービスの入出力に着目しそれぞれのプライバシーコストを計算し評価関数に加える。本研究ではプライバシーデータ項目数を7つ、カテゴリの数は3つとし、それぞれのカテゴリのウェイトを $w_{c1}=0.2$, $w_{c2}=0.4$, $w_{c3}=0.4$ とする。

4.5 評価関数

遺伝的アルゴリズムで用いる評価関数 $F(g)$ は、コスト、レスポンスタイム、稼働率、信頼性、入出力のプライバシーコスト、制約で定義される。本研究の制約 $P'(g)$ は、サービス間の入出力で制約違反があるかないかを求めている。 i を抽象サービス間数、 j をプライバシーデータ数とし以下の式(6)と定義する。

$$P'(g) = \sum_{i=1}^I \sum_{j=1}^J y_{ij}(g) \quad (6)$$

制約が満たしているならば y_{ij} は0、満たしていないなら1とする。本研究では、抽象サービス間数は4なので $I = 4$ 、プライバシーデータ項目数は7なので $J = 7$ である。

$$\begin{aligned} F(g) = & w_{g1}Cost(g) + w_{g2}ResponceTime(g) \\ & + w_{g3}Availability(g) + w_{g4}Reliability(g) \\ & + w_{g5}InPrivacy(g) + w_{g6}OutPrivacy(g) \\ & - w_{g7}P'(g) \end{aligned} \quad (7)$$

$F(g)$ が高いほど、ユーザに適したサービス連携をしていることになる。ここで、 w_{g1}, \dots, w_{g7} はそれぞれの重みであり、全てのウェイトを足すと1になる。また、評価関数の全ての項目において $[0,1]$ に標準化を行う。

5. シミュレーション実験

5.1 実験設定

オンライン上で旅行計画を行うことを考える。旅行計画に必要なサービスのワークフローは5つの抽象サービスがシーケンス構造になっている。飛行機予約サービス、ホテル予約サービス、レンタカー予約サービス、観光・レジャー予約サービス、決済サービスであり、それぞれの抽象サービスを28種類、25種類、26種類、30種類、27種類とする。ワークフローを以下の図4に示す。

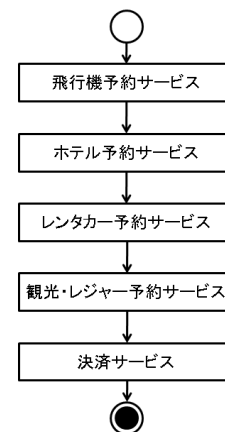


図 4: シミュレーションに用いたワークフロー

また、それぞれのQoSの評価指標の初期値は一様乱数により生成されており、コストは $[50-150]$ 、レスポンスタイムは $[100-200]$ 、稼働率は $[0.9-0.99]$ 、信頼率は $[0.5-1]$ であり、評価のウェイトは $w_{g1}, w_{g2}, w_{g3}, w_{g4}, w_{g5}, w_{g6} = 1$, $w_{g7} = 4$ とした。プロバイダの制約として、制約の強度が異なる制約

表 1: 上位 10 件のサービス

制約 A						制約 B							
順位	サービス					適応度	順位	サービス					適応度
1	[1-21]	[2-1]	[3-9]	[4-5]	[5-24]	1.353	1	[1-19]	[2-24]	[3-9]	[4-1]	[5-5]	1.543
2	[1-19]	[2-18]	[3-9]	[4-5]	[5-24]	1.339	2	[1-19]	[2-24]	[3-9]	[4-9]	[5-0]	1.539
3	[1-21]	[2-1]	[3-16]	[4-5]	[5-9]	1.296	3	[1-26]	[2-1]	[3-12]	[4-12]	[5-0]	1.519
4	[1-19]	[2-1]	[3-16]	[4-17]	[5-5]	1.234	4	[1-26]	[2-24]	[3-9]	[4-1]	[5-23]	1.464
5	[1-19]	[2-1]	[3-9]	[4-5]	[5-24]	1.217	5	[1-19]	[2-1]	[3-12]	[4-12]	[5-0]	1.463
6	[1-21]	[2-14]	[3-7]	[4-5]	[5-24]	1.196	6	[1-19]	[2-24]	[3-25]	[4-12]	[5-0]	1.457
7	[1-19]	[2-1]	[3-16]	[4-5]	[5-24]	1.191	7	[1-26]	[2-1]	[3-25]	[4-12]	[5-0]	1.442
8	[1-19]	[2-1]	[3-9]	[4-5]	[5-5]	1.148	8	[1-26]	[2-24]	[3-25]	[4-8]	[5-0]	1.437
9	[1-21]	[2-1]	[3-7]	[4-5]	[5-24]	1.146	9	[1-19]	[2-1]	[3-12]	[4-8]	[5-9]	1.427
10	[1-21]	[2-1]	[3-7]	[4-17]	[5-5]	1.129	10	[1-19]	[2-1]	[3-12]	[4-8]	[5-5]	1.420

A, 制約 B を設定した. 制約 A は, Hewett ら [Hewett 2009] と同じでサービス間でプライバシーデータの受け渡す時に入出力のデータが同じではないといけな制約とし, 制約 B はデータが同じでなくても同じカテゴリーであれば代替可能である制約とした. 遺伝的アルゴリズムのパラメータは以下の表 2 に示す.

表 2: 遺伝的アルゴリズムのパラメータ

集団の大きさ	交叉率	突然変異率	最大世代数
40	0.7	0.05	100

5.2 実験結果と考察

遺伝的アルゴリズムを用いた上位 10 件のサービスと適応度を制約ごとに表 1 に示す. [1-21] は抽象サービス 1 の候補サービス集合のサービス 21 を表している. 表 1 より制約 A と制約 B ではサービスの入れ替わりや埋もれていたサービスを新たに抽出していることが分かる. 特に抽象サービス 4 では [4-5] から [4-1] や [4-12] に, 抽象サービス 5 では [5-24] から [5-0] や [5-23] に変わっている. これらは, プロバイダのプライバシーポリシーの不一致の影響が大きく, 制約 A で埋もれていたサービスが制約 B で新たに抽出されたと考えられる. 一方, 抽象サービス 1, 2 では制約を変えても同じサービスが上位に多いことが分かる. これらの抽象サービス 1, 2 では, プロバイダのプライバシーポリシーが制約 A で一致していたことが考えられる. これらにより, ユーザのプライバシーポリシーだけではなく, プロバイダのプライバシーポリシーを変えることにより, 新たにサービスが抽出され今まで埋もれていたサービスをユーザがサービス候補から求めているサービスを選択することが可能になると考えられる. また, プロバイダのプライバシーポリシーによって抽象サービスの上位に現れるサービス群が違うことも表 1 で分かる.

6. まとめと今後の課題

本研究では, サービス品質とユーザのプライバシーポリシーに配慮したサービス連携モデルを構築した. その上で提案したモデルの有用性をシミュレーションにより評価した. シミュレーションにより, プロバイダのプライバシーポリシーを変えることでサービスの入れ変わりがおき, またユーザに新たなサービスが抽出された. よってサービス選択において, ユーザのプライバシーポリシーだけではなくプロバイダのプライバシーポリ

シーを考慮することも重要であると考えられる. また, Web サービスを提供するプロバイダは様々なプライバシーポリシーを持っていると考えられ, 今回のシミュレーション結果により様々なプライバシーポリシーを持っているサービスの状況下での新たなサービスの発見には有用だと考えられる.

今後の課題は抽象サービス毎にプロバイダのプライバシーポリシーを変えたり, すべての評価項目を考慮するのではなく, ユーザが要求する評価項目に絞り, それらを用いてサービス選択を行うことも必要と考えられる. また, Web サービスの環境に特化した遺伝的アルゴリズムの改善やマルチエージェントによる大規模なワークフローにおいてもシミュレーションも必要と考えられる.

参考文献

- [青木 2005] 青木利晴: Web サービスコンピューティング, 社団法人電子情報通信学会, (2005).
- [Zeng 2004] L.Zeng, B.Benatallah, A.H.H.Ngu, M.Dumas, J.Kalagnanam, and H.chang: QoS-Aware Middleware for Web Services Composition, *IEEE Transactions on Software Engyneering*, pp311-327, (2004).
- [Canfora 2005] G.Canfora, M.Di Penta: An Approach for QoS-aware Service composition based on genetic Algorithms, *The Genetic and Evolutionary Computation Conference(GECCO'05)*, pp.1069-1075, (2005).
- [Yu 2006] Tao Yu and Kwei-Jay Lin: Modeling and Measuring Privacy Risks in QoS Web Services, *In Proceedings of the The 8th IEEE International Conference on E-Commerce Technology*, (2006).
- [Hewett 2009] Rattikorn Hewett and Phong phun Kijjanayothin: On Securing Privacy in composite Web Service Transactions, *In Proceedings of The 5th International Conference for Internet Technology and secured Transactions(ICITST'09)*, (2009).
- [Xu 2006] Wei Xu, V.N.Venkatkrishnan, R.Sekar and I.V.Ramakrishnan: A Framework for Building Privacy-Conscious Composite Web Services, *IEEE International Conference on Web Services(ICWS'06)*, pp655-662, (2006).