

# 評判情報サイトにおける集合知を用いた製品/サービスの価値推定 - 粒子フィルタを用いたユーザ評価からの価値推定手法の開発 -

高橋 梓      山田和明

東洋大学大学院理工学研究科 東洋大学理工学部機械工学科

s36A01400056@toyo.jp

**概要** 本研究は、粒子フィルタと局所回帰平滑化によって、Amazon.com や食べログなどの評判情報サイトにおけるユーザ評価から製品/サービスの適正な価値を推定する新しい手法を提案する。既存の評判情報サイトでは、製品/サービスに対するユーザ評価を統計処理することで、製品/サービスの適正な価値を推定することが試みられている。しかし、例えばレストランが常に料理やサービスの質を一定に保ったり、全てのユーザが常に料理やサービスの価値を正確に評価することは難しく、ユーザ評価と製品/サービスにはゆらぎが含まれていると考えるのが自然である。そこで本研究では、粒子フィルタと局所回帰平滑化によって、ユーザと製品/サービスに含まれるゆらぎを除去することで、製品/サービスの適正な価値を推定することを試みる。そして、提案手法の有効性を計算機実験を通して検証する。

**キーワード** 評判情報サイト, 集合知, 粒子フィルタ, 自己組織型状態空間モデル, 局所回帰平滑化

## 1 はじめに

Amazon.com<sup>1</sup> や食べログ<sup>2</sup> などのウェブサイトでは、ユーザが利用した製品/サービスに対するレビューや評価を投稿し、共有することができるサービスを提供している。本研究では、このようにユーザのレビューや評価を共有するサービスを提供しているサイトを評判情報サイトと呼ぶこととする。評判情報サイトの魅力は、ミシュランガイドのように訓練された調査員が厳格な基準に従って評価するのではなく、一般ユーザが自身の感じのままに評価するにもかかわらず、多くのユーザが妥当だと思える製品/サービスの適正な価値を推定する点である。ソーシャルネットワークの発展によって、ユーザは自由に情報を発信・共有することで多くの利便性を享受する一方、ネット上には玉石混淆の情報が溢れており、ユーザが個々の情報の価値を直接確かめることは極めて困難になっている [1]。そのため、評判情報サイトのように多くのユーザ評価からコンテンツの適正な価値を推定する技術は今後非常に重要となる。

既存の評判情報サイトでは、多数のユーザ評価を独自のアルゴリズムで処理することでコンテンツの適正な価値を推定する試みが行われている。しかし、例えばレストランの場合、料理やスタッフの質が一定でも多くの客が一度に来店すると配膳が遅れるなど、外的な要因によりサービスの質が低下する。また、レストランのスタッフが経験を積むことで料理や接客の質が向上することもある。一方、人はコンテンツの質が同じでも時間が経つと評価に一貫性を保てなくなるなど、評価がゆらぐこと

が知られている [2]。また、人は同じ料理でも空腹時と満腹時では評価が異なり、同じワインでも高い価格の方により高い満足感を得ることが報告されている [3]。このように、コンテンツの質やユーザの評価は様々なゆらぎによって、時々刻々と変化すると考えるのが自然である。

我々の研究グループでは、コンテンツの適正な価値を推定するには、コンテンツに発生するゆらぎと、ユーザが評価する際に発生するゆらぎ、の両方を考慮する必要があると考えている。また、コンテンツとユーザ評価に発生するゆらぎの特性は事前には知ることができないため、ユーザがコンテンツを評価したインタラクションの履歴からゆらぎの特性を推定する必要がある。そこで本研究では、粒子フィルタの一種である粒子スムーザ (Particle Smoother: PS) [4] と局所回帰平滑化の一種である Loess [5] によって、コンテンツの適正な価値を推定する。また、粒子スムーザに自己組織型状態空間モデル [9, 10] を導入することで、コンテンツおよびユーザ評価に発生するゆらぎの特性もあわせて推定する手法を提案する。

本稿では、次章において関連研究について説明し、次に、評判情報サイトのモデル化を行う。そして、提案手法である評判情報システムについて説明し、計算機実験を通して提案手法の有効性を検証する。

## 2 関連研究

本章では、これまでに研究されてきたユーザ評価からコンテンツの価値を推定する方法や既存のサービスについて説明し、提案手法と従来手法の違いについて述べる。まず、eBay<sup>3</sup> などのオークションサイトでは、売

Copyright is held by the author(s).

The article has been published without reviewing.

<sup>1</sup>www.amazon.com

<sup>2</sup>tabelog.com

<sup>3</sup>www.ebay.com

り手と買い手が互いに相手の取引を評価し、その結果を合計することで各ユーザの信頼度を算出している。しかし、ユーザ評価はゆらぎを含むため、単純なユーザ評価の合計では適切な信頼度を得られない。そのため酒井ら [6] は、ユーザが互いに評価した値から確率的近似法を用いて適切なユーザの信頼度を推定する方法を提案している。そして、評価のゆらぎが大きいユーザや嘘の評価を付けるユーザが存在しても提案手法が頑健であることをシミュレーションにより示している。また、小林ら [7] は、互いに評価したユーザ間のネットワーク関係を隣接行列で記述し、HITS アルゴリズム [8] によって各ユーザの信頼度を推定する方法を提案している。そして、Yahoo!オークションの実データを用いて提案手法の有効性を示している。

一方、評判情報サイトはオークションサイトのようにユーザ間で互いに評価しないため、上記の手法をそのまま利用することはできない。そのため、食べログやアットコスメ<sup>4</sup>などの評判情報サイトでは、レビュアーの投稿数や他のユーザがレビューを評価した数など様々な情報からそのレビュアーの信頼度を算出し、レビュアーが付けたコンテンツの評価値を信頼度で重み付けして合計することでコンテンツの価値を推定している。しかし、この手法ではコンテンツやユーザ評価に発生するゆらぎを考慮していない。

そこで本研究では、粒子フィルタの一種である粒子スムーザと局所回帰平滑化手法の一種である Loess によって、コンテンツやユーザ評価に発生するゆらぎを除去することでコンテンツの適正な価値を推定する。また、コンテンツやユーザ評価に発生するゆらぎの特性を予め知ることできないため、粒子スムーザに自己組織型状態空間モデルを導入することで、コンテンツやユーザ評価に発生するゆらぎの特性を同時に推定する。

### 3 評判情報サイトのモデル化

本章では、評判情報サイトにおいて、ユーザが利用したコンテンツを評価する過程をモデル化する。まず、評判情報サイトには  $m$  個のコンテンツが登録されており、 $n$  人のユーザがコンテンツを利用して評価するものとする。図 1 に示すように、 $j$  番目のコンテンツ  $C_j$  の時刻  $t$  における適正な価値を  $x_t$  とする。しかし、ユーザ  $U_i$  はコンテンツ  $C_j$  を利用したときコンテンツに発生するゆらぎ  $v_t$  によって、式 (1-a) のように適正な価値とは異なる  $x'_t$  を観測する。また、ユーザはコンテンツを評価する際に発生するゆらぎ  $w_t$  によって、式 (1-b) のようにコンテンツを  $y_t$  と評価する。したがって、ユーザ  $U_i$  はコンテンツ  $C_j$  を利用したとき、コンテンツに発生す

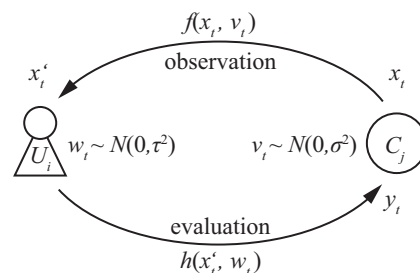


図 1 評判情報サイトにおけるユーザ評価プロセス。

るゆらぎとユーザが評価する際に発生するゆらぎによって、コンテンツ  $C_j$  の適正な価値  $x_t$  からずれた評価値  $y_t$  を付ける。

$$x'_t = f(x_t, v_t), \quad v_t \sim N(0, \sigma^2) \quad (1-a)$$

$$y_t = h(x'_t, w_t), \quad w_t \sim N(0, \tau^2) \quad (1-b)$$

ただし、 $f$  と  $h$  は状態変数  $x$  に関する非線形関数であり、コンテンツのゆらぎ  $v_t$  は平均ゼロ、分散  $\sigma^2$  の白色雑音  $N(0, \sigma^2)$ 、ユーザの評価のゆらぎ  $w_t$  は平均ゼロ、分散  $\tau^2$  の白色雑音  $N(0, \tau^2)$  とする。粒子フィルタでは、式 (1-a) をシステムモデル、式 (1-b) を観測モデルと呼び、 $v_t$  および  $w_t$  をそれぞれシステムノイズ、観測ノイズと呼ぶ。なお、システムノイズは、システムをモデル化した際に発生するズレを表す。

## 4 評判情報システム

本章では、提案する評判情報システムの概要を説明し、評判情報システムで用いる粒子フィルタと局所回帰平滑化について説明する。なお、提案手法で用いる粒子フィルタは、推定精度を上げるために固定ラグ平滑化 [4] を導入した粒子スムーザ (Particle Smoother: PS) を用いる。また、システムノイズと観測ノイズを推定するために、PS に自己組織型状態空間モデルを導入している。固定ラグ平滑化と自己組織型状態空間モデルは、粒子フィルタの状態ベクトルを拡張することで容易に導入できる。そのため本章では、まず、粒子フィルタと局所回帰平滑化を用いた評判情報システムのアルゴリズムを説明し、その後、粒子フィルタを固定ラグ平滑化と自己組織型状態空間モデルへ拡張する方法を説明する。

### 4.1 システムの概要

評判情報システムは、ユーザやコンテンツが登録されるとユーザモデル  $U$  とコンテンツモデル  $C$  を作成する。ユーザモデル  $U$  は  $N$  個の粒子を持ち、各粒子はユーザの評価のゆらぎ  $w$  の大きさを表す分散  $\tau^2$  を記憶している。コンテンツモデル  $C$  も  $N$  個の粒子を持ち、各粒子はコンテンツの評価値  $x$  とコンテンツに発生するゆらぎ  $v$  の大きさを表す分散  $\sigma^2$  を記憶している。そして、

<sup>4</sup>www.cosme.net

ユーザがコンテンツを利用したとき，以下の手順に従ってコンテンツの適正な評価値と，同時にユーザとコンテンツに発生するゆらぎを推定する．

- (1) ユーザ  $i$  がコンテンツ  $j$  を利用して評価すると，評判情報システムはユーザモデル  $U_i$  とコンテンツモデル  $C_j$  が記憶している変数を結合した状態ベクトルを生成する．そして，生成された状態ベクトルから構成された粒子集団  $\{z_{t-1|t-1}^{(i)}\}_{i=1}^N$  を状態空間に配置する．時刻  $t-1$  におけるコンテンツの評価値の確率分布はこの粒子集団によって近似される．
- (2) ユーザが付けた評価値  $y_t$  に基づいて粒子フィルタによって時刻  $t$  における粒子集団  $\{z_{t|t}^{(i)}\}_{i=1}^N$  を生成する．
- (3) 状態ベクトルの要素を分離し，ユーザモデルとコンテンツモデルの変数を更新する．
- (4) 時刻  $t$  の粒子集団からコンテンツの評価値  $\hat{x}_{t|t}$  を計算する．
- (5) Loess によってコンテンツの評価値  $\bar{x}_{t|t}$  を計算する．

提案手法では，粒子フィルタの推定値  $\hat{x}_t$  を Loess により平滑化した値  $\bar{x}_t$  をコンテンツの評価値とする．これは，Loess により粒子フィルタの推定値に含まれるコンテンツのノイズを除去するためである．

#### 4.2 粒子フィルタ

粒子フィルタは図2に示すように，まず，多数の粒子をシステムモデルに基づいて状態空間に散布し，システムの状態を近似する予測分布を生成する．次に，各粒子を観測モデルで写像した値と観測値  $y_t$  を比較し，各粒子の値がシステムの状態  $x_t$  にどれだけ当てはまるかを尤度により評価する．そして，尤度に基づいて粒子をリサンプリング（復元抽出）し，フィルタ分布を生成する．このとき，システムの状態はフィルタ分布の平均値として求められる．粒子フィルタは上記操作を繰返すことで観測値からシステムの状態を推定する．以下に粒子フィルタのアルゴリズムと，固定ラグ平滑化および自己組織型状態空間モデルの導入方法について説明する．

##### 4.2.1 アルゴリズム

粒子フィルタのアルゴリズムは以下ようになる．

- (1) 初期分布を近似する粒子集団  $\{x_{0|0}^{(i)}\}_{i=1}^N$  ( $x_{0|0}^{(i)} \sim p_0(x)$ ) を生成する．ただし， $p_0(x)$  は  $x$  の時刻  $t=0$  における初期分布を表す．
- (2)  $t=1, \dots, T$  について (a) ~ (c) のステップを実行する．

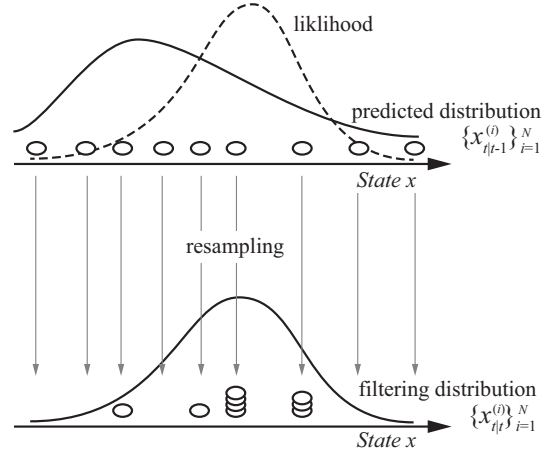


図2 粒子フィルタの概念図.

##### (a) 尤度計算

各  $i$  ( $i=1, \dots, N$ ) について (i) ~ (iiii) を実行する．

- (i) 乱数  $v_t^{(i)} \sim p(v)$  を生成する．
- (ii)  $x_{t|t-1}^{(i)} = f_t(x_{t-1|t-1}^{(i)}, v_t^{(i)})$  を計算し，予測分布の粒子集団  $\{x_{t|t-1}^{(i)}\}_{i=1}^N$  を得る．
- (iii)  $\lambda_t^{(i)} = p(y_t|x_{t|t-1}^{(i)})$  を計算する．
- (iiii)  $\beta_t^{(i)} = \lambda_t^{(i)} / \sum_{i=1}^N \lambda_t^{(j)}$  を求める．

##### (b) リサンプリング

粒子集団  $\{x_{t|t-1}^{(i)}\}_{i=1}^N$  の中から各粒子  $x_{t|t-1}^{(i)}$  を  $\beta_t^{(i)}$  の確率で重複を許して復元抽出し，新たな粒子集団  $\{x_{t|t-1}^{(i)}\}_{i=1}^N$  を生成する．

##### (c) 時刻 $t$ の状態推定

得られた  $N$  個のサンプルからフィルタ分布を近似する  $\hat{x}_{t|t} = \frac{1}{N} \sum_{i=1}^N x_{t|t-1}^{(i)}$  を構成する．

なお，観測ノイズ  $w_t$  が平均ゼロ，分散  $\tau^2$  の正規分布に従う場合，粒子  $x_{t|t-1}^{(i)}$  の尤度  $p(y_t|x_{t|t-1}^{(i)})$  は次式で求められる．

$$p(y_t|x_{t|t-1}^{(i)}) = \frac{1}{\sqrt{2\pi\tau^2}} \exp \left[ -\frac{(y_t - H(x_{t|t-1}^{(i)}))^2}{2\tau^2} \right] \quad (2)$$

##### 4.2.2 固定ラグ平滑化

固定ラグ平滑化 [4] は，状態ベクトル  $z_{t|t-1}^{(i)}$  を式 (3) のように拡張し，この拡張状態ベクトルに対して粒子フィルタと同様にリサンプリングすることで実現できる．最後にリサンプリングにより得られた粒子集団  $\{z_{t|t}^{(i)}\}_{i=1}^N = \{[z_{t|t}^{(i)}, \dots, z_{t-L|t}^{(i)}]'\}_{i=1}^N$  から  $\{z_{t-L|t}^{(i)}\}_{i=1}^N$  の部分を取り出すことで，時刻  $t-L$  から時刻  $t$  の情報に基づいて平滑

化した時刻  $t - L$  の状態を求めることができる．

$$\hat{z}_{t|t-1}^{(i)} = [z_{t|t-1}^{(i)}, z_{t-1|t-1}^{(i)}, \dots, z_{t-L|t-1}^{(i)}]' \quad (3)$$

#### 4.2.3 自己組織型状態空間モデル

自己組織型状態空間モデルは、式 (4) に示すように、状態  $x_t$  に超パラメータ  $\theta_t$  を含めた状態ベクトル  $z_t$  を生成する．ただし、超パラメータ  $\theta_t$  はシステムノイズ  $v_t$  の分散  $\sigma_t^2$  と観測ノイズ  $w_t$  の分散  $\tau_t^2$  から構成されている．生成した状態ベクトル  $z_t$  に対して粒子フィルタと同様の操作を繰返すことで、システムノイズと観測ノイズの分散を同時に推定することができる<sup>5</sup>．このとき、システムモデルと観測モデルは式 (5) と記述でき、非線形関数  $F$  と  $H$  は式 (6) となる．

$$z_t = \begin{bmatrix} x_t \\ \theta_t \end{bmatrix}, \quad \theta_t = \begin{bmatrix} \log_{10} \sigma_t^2 \\ \log_{10} \tau_t^2 \end{bmatrix} \quad (4)$$

$$z_t = F(z_{t-1}, v_t) \quad (5-a)$$

$$y_t = H(z_t, w_t) \quad (5-b)$$

$$F(z_{t-1}, v_t) = \begin{bmatrix} f(x_{t-1}, v_t) \\ \theta_{t-1} + \epsilon_t \end{bmatrix} \quad (6-a)$$

$$H(z_t, w_t) = h(z_t, w_t) \quad (6-b)$$

ただし、式 (6-a) の  $\theta_t = \theta_{t-1} + \epsilon_t$  は、超パラメータ  $\theta_t$  の時間変化を表し、 $\epsilon_t = [\zeta_t, \eta_t]'$  はそれぞれ  $\zeta_t \sim N(0, \nu^2)$ 、 $\eta_t \sim N(0, \xi^2)$  とする． $\nu$  と  $\xi$  は超パラメータを特徴付けるパラメータであるため、超々パラメータと呼ばれる．

#### 4.3 局所回帰平滑化

局所回帰平滑化の一種である Loess は式 (7) に基づいて重み付き線形最小二乗回帰を実行することで行われる．このとき回帰重み  $w(x_t)$  は決められた範囲内の各データ点に対して式 (8) から求められる．回帰重みはデータ  $x_t$  において最大となり、そこから離れるにつれて小さくなる．したがって、回帰重みはデータ  $x_t$  に対して左右対称となるが、データの端点（始点と終点）では非対称となる．通常、Loess は予め用意されたデータに対して実行されるが、本稿では新しいデータが観測される度に Loess を実行することで逐次データを平滑化する．ただし、データの端点では回帰重みが非対称となり、端点の重みが最大となるため、本稿では、観測した最新のデータからラグタイム分だけ過去のデータを利用する．

$$\sum_{t=1}^N w(x_t)(y_t - a + bx_t - cx_t^2)^2 \quad (7)$$

<sup>5</sup> $\sigma^2$  と  $\tau^2$  は正値性を保つため対数値を用いる．

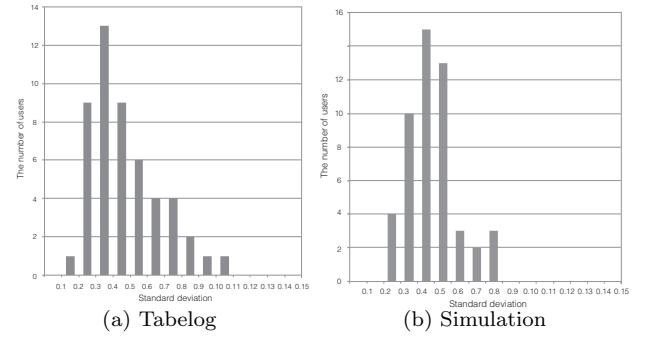


図3 ユーザ評価のゆらぎの大きさの分布.

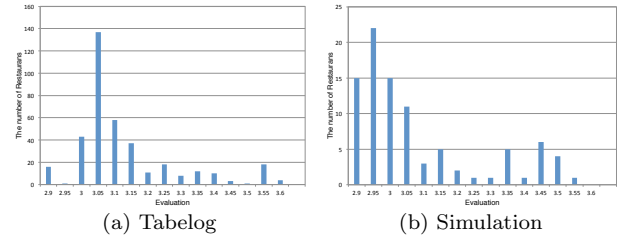


図4 食べログに掲載されているレストランの評価値の分布.

$$w(x_t) = \left(1 - \left\| \frac{x_i - x_t}{d(x)} \right\|^3\right)^3 \quad (8)$$

ただし、 $x_t$  は平滑化する時刻  $t$  のデータであり、 $x_i$  は範囲内のデータを表す． $d(x)$  は範囲内で  $x_t$  から最も離れたデータの距離を表す．

## 5 計算機実験

### 5.1 実験設定

#### 5.1.1 ユーザのモデル化

ユーザ評価のゆらぎをモデル化するために、既存の評判情報サイトである食べログのユーザ評価を分析する．まず、食べログから投稿の多いユーザを選び、各ユーザが利用したレストランからユーザ評価数の多いものだけを選出する．そして、各レストランのユーザ評価の平均を計算し、そのレストランの価値とする．次に、各レストランの価値と選ばれたユーザが付けた評価値との差からユーザ評価のゆらぎの大きさ（標準偏差）を算出する．図 3(a) は食べログから選んだユーザ 50 名の標準偏差のヒストグラムである．なお、グラフの横軸は標準偏差であり、縦軸はユーザ数を表す．この結果から、ほとんどのユーザは評価のゆらぎが小さく、一部のユーザのゆらぎが大きいことがわかる．実験ではユーザ評価のゆらぎの大きさ（標準偏差）は対数正規分布に従うものとし、棄却法によりユーザの標準偏差を図 3(b) のように決定する．

5.1.2 コンテンツのモデル化

レストランの適正な評価値やゆらぎを予め測定することは困難である．そのため、実験では各レストランの評価値は常に一定とし、レストランのゆらぎとして平均ゼロ、分散 0.04 の白色雑音を用いる．ただし、各レストランの評価値は、食べログに掲載されているレストランの評価値を分析した結果、図 4(a) のように混合正規分布に従うため、棄却法により図 4(b) のように決定する．

5.1.3 ユーザ評価プロセスのモデル化

実験では、自己組織型状態空間モデルを導入した粒子スモージにより各ユーザの評価のゆらぎを推定し、その後、評価のゆらぎを推定されたユーザが、新規レストランを評価したとき、提案手法の推定精度が向上するか確認する．そのため、まず、ユーザ 50 名にレストラン 50 軒を評価してもらい、その後、新規レストラン 50 軒を評価してもらう．この一連の流れを 1 トライアルとし、実験では 25 トライアル行う．実験では、ユーザとレストランが記憶する粒子数を  $N = 500$  とする．また、自己組織型状態空間モデルによりユーザ評価のゆらぎを推定するために、粒子の初期分布を  $U([0.0, 0.25])$  の一様乱数により決定し、超々パラメータを  $\xi = 0.05$  とする．一方、レストランの評価値とレストランに発生するゆらぎを推定するために、粒子の初期分布をそれぞれ  $U([0.0, 5.0])$ 、 $U([0.0, 0.04])$  の一様乱数により決定し、超々パラメータを  $\nu = 0.02$  とする．固定ラグ平滑化と Loess のラグタイムをそれぞれ  $L_{ps} = 5$ 、 $L_{loess} = 5$  とする．

5.2 実験結果

図 5 はあるレストランの価値を粒子スモージ (PS) で推定した結果と提案手法 (PS と Loess) の推定結果を表す．グラフの横軸は日数であり、縦軸はレストランの評価値を表す．赤線はレストランの適正な価値であり、青の \* は個々のユーザ評価を表す．緑、ピンクの線はそれぞれ PS、提案手法 (PS と Loess) の推定結果を表す．なお、提案手法は PS と Loess のラグタイムを 5 と設定しているため、10 日前の推定値を出力する．そのため、グラフは 40 日までの結果がプロットされている．この結果から、PS はユーザがレストランの適正な評価値から外れた評価を付けても補正していることがわかる．また、提案手法は PS の推定結果に Loess を掛けることで、PS のみで推定するより推定精度が向上している．

次に、提案手法によって、あるユーザ評価のゆらぎの大きさ (標準偏差) を推定した結果を図 6 に示す．グラフの横軸は日数、縦軸はユーザ評価のゆらぎの大きさ (標準偏差) を表す．赤と緑の線はそれぞれユーザ評価の標準偏差と自己組織型状態空間モデルによる推定結果を表す．この結果から、提案手法が 40 日頃にはユーザ評価の標準偏差を精度よく推定していることがわかる．

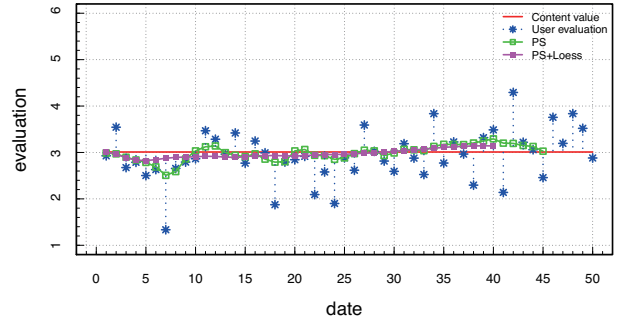


図 5 あるレストランの価値の推定結果.

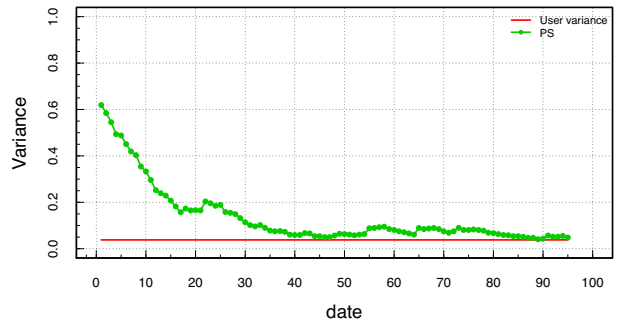
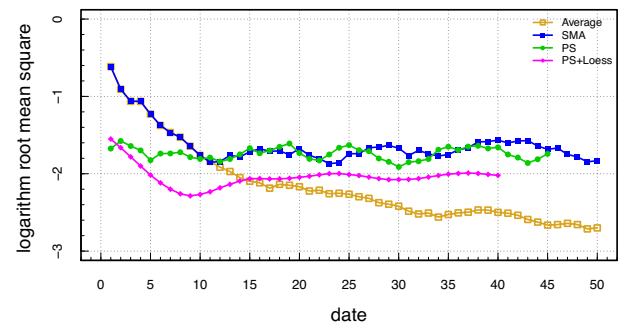
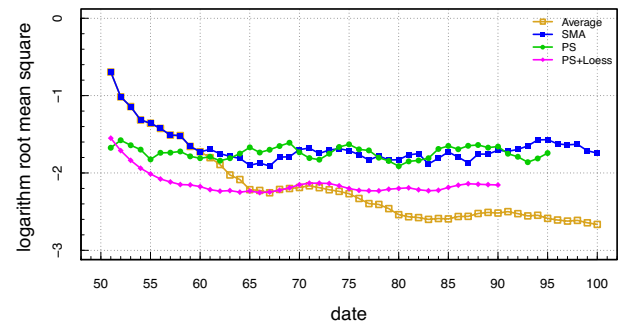


図 6 あるユーザ評価のゆらぎの標準偏差.



(a) 1 ~ 50 日の推定結果



(b) 51 ~ 100 日の推定結果

図 7 レストランの価値と推定値との二乗平均平方根.

図 7 はユーザ 50 名が 1 ~ 50 日の間にレストラン 50 軒を評価した後 (図 7(a)), 51 ~ 100 日の間に新規レストラン 50 軒を評価した場合 (図 7(b)) のレストランの価値と各手法の推定値との二乗平均平方根の対数を示

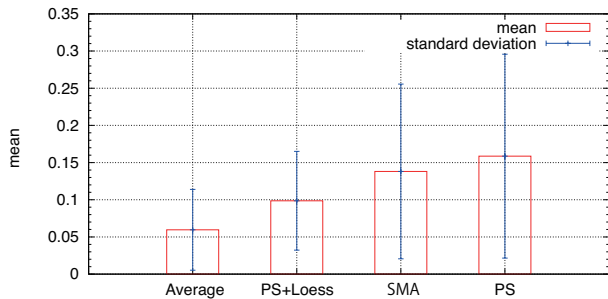


図8 レストランの価値と推定値との誤差の平均と標準偏差。

す。なお、推定手法として平均 (Average), 単純移動平均 (SMA), PS, 提案手法 (PS と Loess) を用い、グラフではそれぞれ橙, 青, 緑, ピンクの線によって表す。図 7(a) と 7(b) の結果から、提案手法 (PS と Loess) は 15 日頃までは他の手法より推定精度が高いことがわかる。これは提案手法が、PS によりユーザ評価を補正し、Loess により平滑化しているためだと考えられる。図 7(b) の結果から、提案手法の推定精度は自己組織型状態空間モデルを用いてユーザ評価の標準偏差を推定することで、1~50 日の結果 (図 7(a)) より向上していることがわかる。そのため、提案手法は他の手法より少ないユーザ評価から精度よくレストランの適正な価値を推定できると期待できる。

各手法の推定精度の差を詳細に分析するために、図 8 に 90 日目の個々のレストランの価値と各手法の推定値との平均誤差<sup>6</sup>と標準偏差を示す。90 日目の提案手法と単純平均との平均誤差の差は約 0.05 であり、標準偏差はほぼ同じであった。今回の実験では、レストランの価値が常に一定の場合を想定しているため、単純平均の推定値の方が良い結果となった。しかし、実際のレストランの価値は料理、スタッフ、値段など様々な要因によって変化すると考えられる。そのため、今後、提案手法の推定精度を向上させるとともに、各コンテンツの価値が変動する場合の提案手法の有効性を検証する必要がある。

## 6 おわりに

本稿では、評判情報サイトに投稿されたユーザ評価から製品/サービスの適正な価値を、粒子フィルタと局所回帰平滑化を用いることで推定する新しい手法を提案した。提案手法は、粒子フィルタに自己組織型状態空間モデルと固定ラグ平滑化を導入することで、製品/サービスの価値と同時に、ユーザと製品/サービスに発生するゆらぎも精度よく推定できるという特徴を持つ。

提案手法の有効性を検証するために、既存のレストラン評判情報サイトを参考にシミュレーションを構築し、

計算実験を行った。実験の結果、提案手法は粒子フィルタによってユーザ評価のゆらぎを低減し、粒子フィルタの推定結果を Loess で平滑化することで、粒子フィルタ単体より推定精度が向上することを確認した。また、自己組織型状態空間モデルを用いてユーザ評価のゆらぎの大きさを推定することで、ユーザ評価のゆらぎが未知の場合に比べ、ユーザ評価数が少ない段階から推定精度が向上することを確認した。

今後の課題として、粒子フィルタの推定精度の向上と、ユーザ数とコンテンツ数を増やした場合の提案手法の推定精度を検証し、実問題に適用する予定である。

## 謝辞

本研究の一部は、JSPS 科研費 25730185、および、JST RISTEX 問題解決型サービス科学研究開発プログラム採択プロジェクト「価値創成クラスモデルによるサービスシステムの類型化とメカニズム設計理論の構築」の助成を受けたものです。

## 参考文献

- [1] 武田英明, 信頼の Web, コンピュータソフトウェア, Vol.22, No.4, pp.19-25, 2005.
- [2] 神蔭敏弘, 協調フィルタリングの課題: プライバシー, サクラ攻撃, 評価値のゆらぎ, 情報処理, Vol.48, No.9, pp.966-971, 2007.
- [3] レナード ムロディナウ, 田中三彦 (訳): たまたま - 日常に潜む「偶然」を科学する, ダイヤモンド社, 2009.
- [4] 樋口知之: データ同化入門 (予測と発見の科学), 朝倉書店, 2011.
- [5] William S. Cleveland and Susan J. Devlin., Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, Journal of the American Statistical Association, Vol.83, No.403, pp.596-610, (1988).
- [6] 酒井隆道, 寺田賢二, 樺肅之, 確率的近似法を用いた頑強なオンライン評判メカニズム, 電子情報通信学会論文誌, J88-D-I(5), pp.958-968, 2005.
- [7] 小林真雄, 安藤哲志, 伊藤孝行, Auction Network Trust: 電子商取引ネットワークにおけるユーザ間の関係を利用した評判メカニズム, 電子情報通信学会論文誌, J92-D(11), pp.1869-1880, 2009.
- [8] L. Li, Y. Shang, and W. Zhang, Improvement of hits-based algorithms on web documents, Proceedings of WWW2002, pp.527-535, 2002.
- [9] 北川源四郎, 樋口知之, 知識発見と自己組織型の統計モデル, bit 別冊「発見科学とデータマイニング」, pp.159-168, 2000.
- [10] 市村直幸, 自己組織型状態空間モデルを用いた運動軌跡のフィルタリング, 情報処理学会誌: コンピュータビジョンとイメージメディア, Vol.43, No.SIG 11(CVIM 5), pp.92-104, 2002.

<sup>6</sup>平均誤差はレストラン 50 軒の 25 トライアル分のデータから算出している。