

機関番号：82626

研究種目：若手研究（B）

研究期間：2009～2010

課題番号：21700179

研究課題名（和文） 小売りサービスにおける大規模データと熟練従事者の経験を統合した  
計算モデルの研究研究課題名（英文） Research of computational customer modeling based on large-scale  
datasets and various situations in retail service

研究代表者

石垣 司 (ISHIGAKI TSUKASA)

独立行政法人産業技術総合研究所・サービス工学研究センター・産総研特別研究員

研究者番号：20469597

研究成果の概要（和文）：本研究は小売サービスの現場で観測される大規模な購買履歴データと熟練サービス従事者の経験を統合した計算モデルを構築することを目的とした。デパートにおける大規模なID-POSデータ（1年間分の約3,000,000件の購買履歴情報）を取得し、顧客行動計算モデル化のため、顧客と商品のセグメント化を確率的な潜在クラス分析法を導入することで実行した。また、サービス従事者の経験として天気や曜日などの状況依存性を考慮するための情報も計算モデル化のための変数として利用した。顧客属性、顧客と商品の自動セグメント、状況依存変数からその確率的構造をベイジアンネットワークにより計算モデル化し、その構造から顧客行動に関する知見の抽出を行った。また、推定されたベイジアンネットワークの構造から、購買行動は状況依存変数から大きな影響を受けることが分かった。本研究成果はデパート以外の小売サービス業で観測・保存されているID-POSデータに対しても水平展開や発展的応用が可能な技術である。

研究成果の概要（英文）：This research proposed an actual service support system using discovery of category-based customer behavior knowledge. The method is realized by modeling a customers' purchase behavior with some purchase situations or conditions using massive point of sales data with ID-POS data in a department store chain. We automatically generate categories of customers and items based on a purchase patterns identified in ID-POS data using probabilistic latent semantics indexing. We produce a Bayesian network model including the customer and item categories, situations and conditions of purchases, and the properties and demographic information of customers. This method is applicable for marketing support, service modeling, and decision making in various business fields, including retail services.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,000,000	300,000	1,300,000
2010年度	600,000	180,000	780,000
年度			
年度			
年度			
総計	1,600,000	480,000	2,080,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：サービス工学、大規模データ、ID付POSデータ、消費者行動モデル、確率的意

## 味解析、ベイジアンネットワーク、

### 1. 研究開始当初の背景

近年、人間の日常生活現場がセンサネットワーク化し、多くの企業で ID 付き POS データ、電子マネーなどの日常生活の行動履歴を記録した大規模データが蓄積されるようになった。これらの大規模データは、その分析により顧客の行動を理解しマーケティングへつなげるための有力な材料として多くの企業に注目されている。情報量が爆発的に増大し続ける大規模実データは人間の日常購買行動により発生するため、そのデータは人間行動の非線形性、人間の状態に影響を与える交互作用の影響を大きく受けている。しかし、古典的な統計手法のほとんどでは線形独立性を強く仮定しており、非線形性や交互作用の取り扱いが困難である。そのため、多くの企業では大規模データを収集するためのシステムインフラの開発と維持のみにコストを費やし、集めたデータの分析・活用という点では十分な費用対効果が得られていないことが申請者の調査により明らかになってきている。

顧客の消費行動のモデル化の試みは古くから盛んに行われてきた。それらは普遍的な消費者の購買意思決定の原理を追求することを主目的としている。しかしながら、消費者の意思決定は合理的とは限らず、またグローバル化・ネットワーク化により消費者がもつ価値観の予測が難しくなったこともあり、それらの決定論的枠組みによる顧客行動モデルでは不確実性を含む人間のモデル化には限界があった。加えて、要請される技術と対象が大量生産アプローチから顧客の小規模セグメントへと、さらには個人を対象とした one-to-one サービスへと遷移しているため、それらを包括しかつ実際の現場で望まれている大規模な購買履歴データの実サービス現場へのフィードバックを図るためには、計算機の導入は不可欠となってきている。従来の消費者行動理論で作成されるモデルは消費者の意思決定の要因までも記号的に扱っていたため、計算機による演算を必要とする大規模実データとの融和性は非常に低かった。そのため、従来の消費者行動モデルでは現在のサービス産業が必要としている小規模セグメントや個人を対象としたマーケティングには適しているとは言えない。

一方、近年の統計科学、機械学習、データマイニング技術などの発展により、従来の解析手法では難しかった、複雑な人間の購買行動の解析には必須となる非線形性や交互作用の取り扱いが大規模データにおいても可能となってきている。また、ベイジアンネットワークに代表されるグラフ構造を利用した確率論的計算モデルであるグラフィカルモデル

とその効率的な確率推論法も意欲的に研究されており、大規模実データと計算論的モデルの親和性は高くなってきている。

ベイジアンネットワークでは従来の線形的・決定論的モデルでは困難であった不確実性、交互作用、非線形性の問題に対し、非決定論的な確率グラフ構造で交互作用を記述するとともにその非線形性は離散的な条件付き確率分布として表現する枠組みを提供している。しかし現在まで、小売りサービスにおける大規模実データとベイジアンネットワークの統合という観点からの研究はほとんど行われていない。

### 2. 研究の目的

本研究では、「熟練サービス従事者の経験的知識と大規模データを取り込んだ、小売りサービスでの予測・制御を目的とした顧客行動計算モデル」に関する研究を行う。そこでは非線形性や交互作用を含む非決定論的な確率モデルを含むベイジアンネットワークを用いる。ベイジアンネットワークでは、モデル対象の交互作用を非決定論的な確率グラフ構造で記述し、その非線形性は離散的な条件付き確率分布として表現するため、従来は困難であった小売りサービスにおける顧客行動のモデル化を各種の困難を排した形式で実行できる可能性が大きい。また、ベイジアンネットワークによる顧客行動モデルへ熟練サービス従事者のもつ経験と知識を導入することで、顧客行動予測のためのシミュレートが可能となることを明らかにする。

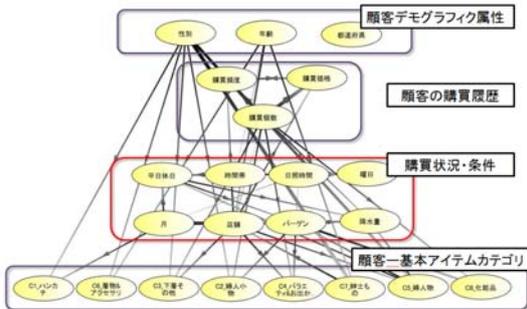
### 3. 研究の方法

本研究は以下の方法によって実行された。

- (1) 大規模データの取得  
平成 21 年度データ解析コンペティションにより提供された百貨店の ID-POS データを取得した。約 3 百万件の購買履歴データを使用可能となった。
- (2) 大規模データの変換、整理。  
(1)により取得した大規模データを目的に即したデータ形式に変換し、天気や顧客属性などの経験・知識につながる状況変数の情報を付与した。
- (3) 確率的潜在意味解析の実装と実行  
(2)のデータに対し、顧客と商品のカテゴリ化を行うための確率的潜在意味解析のプログラムを実装した。また、実際に顧客と商品のカテゴリ抽出を行った。
- (4) (2)で付与した状況変数と(3)で抽出したカテゴリを(1)で取得した購買履歴データに付与し、ベイジアンネットワークモデルの構築を行った。
- (5) (4)で作成したベイジアンネットワークモデル

を利用して、実務のための知識発見を行った。

- (6) より大規模なデータ数、顧客数、商品数に対する本手法の適用の限界を探るため、64bitマシンによるカテゴリ生成とベイジアンネットワークモデル構築のための実験を行った。
- (7) 本研究により達成された成果の応用可能な社会的ニーズの探索を行った。



図：作成したベイジアンネットワークモデル

#### 4. 研究成果

本研究は小売サービスの現場で観測される大規模な購買履歴データと熟練サービス従事者の経験を統合した計算モデルを構築することを目的としている。ここでは小売サービスの中でもデパートを対象に研究を推進した。デパートの購買履歴データである大規模な ID-POS データ（1年間分の約 3,000,000 件の購買履歴情報）を取得した。このデータを用いて顧客行動の計算モデル化を行うために顧客と商品のセグメント化を確率的な潜在クラス分析法を導入することで実行した。このセグメント化技術では適度に購買・被購買パターンの情報を保持したまま、顧客と商品の分類を自動的に発見できることを確認した。また、サービス従事者の経験として天気や曜日などの状況依存性を考慮するための情報も計算モデル化のための変数として利用した。顧客のデモグラフィック属性、顧客と商品の自動セグメント、状況依存変数からその確率的構造をベイジアンネットワークにより計算モデル化し、その構造から顧客行動に関する知見の抽出を行った。また、推定されたベイジアンネットワークの構造から、購買行動は状況依存変数から大きな影響を受けることが分かった。

また、より大規模な顧客数、商品数での顧客行動のモデル化を可能にするための研究を主に行った。そこでは、64bit の計算機環境において十分な量のメモリを使用することができれば 10,000 顧客、10,000 商品程度の組み合わせに対しても、実業務に使用可能な程度の計算負荷で顧客行動の計算モデルを構築できることを確認した。以上の研究成果はデパート以外の小売サービス業で観測・

保存されている ID-POS データに対しても水平展開や発展的応用が可能な技術である。本研究の成果は平成 21 年度データ解析コンペティション日本マーケティングサイエンス学会 ID 付き POS データ活用部会において発表し優勝を獲得した。

#### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 1 件）

- ① 石垣司、竹中毅、本村陽一、百貨店 ID 付き POS データからのカテゴリ別状況依存変数間関係の自動抽出法、オペレーションズ・リサーチ, Vol. 56, No. 2, pp. 77-83, 2011(査読有)

〔学会発表〕（計 6 件）

- ① 石垣司、大規模データを活用した生活者起点のサービス工学へ向けて、応用統計計量ワークショップ, 3 月 9 日, 仙台, 2011(依頼講演)
- ② T. Ishigaki, T. Takenaka and Y. Motomura, Customer-Item Category Based Knowledge Discovery Support System and Its Application to Department Store Service, IEEE Asia-Pacific Services Computing Conference (IEEE APSCC 2010), pp. 371-377, December 6-10, in Hangzhou, China. 2010
- ③ 石垣司、本村陽一、竹中毅、陳希、緒方大樹, PLSA とベイジアンネットワークによる百貨店 ID-POS データからの知識発見ツール, 経営科学系研究部会連合協議会平成 21 年度データ解析コンペティション成果報告会, 3 月 26 日, 東京, 2010
- ④ 石垣司、本村陽一、竹中毅、陳希、緒方大樹, PLSA とベイジアンネットワークによる百貨店 ID-POS データからの知識発見ツール, 経営科学系研究部会連合協議会平成 21 年度データ解析コンペティション最終発表会, 3 月 13 日, 東京, 2010
- ⑤ 石垣司、本村陽一、竹中毅、陳希、緒方大樹, 顧客-基本アイテムセグメントと購買状況・行動の確率的構造の発見法, 経営科学系研究部会連合協議会平成 21 年度データ解析コンペティション中間発表, 2 月 6 日, 東京, 2010
- ⑥ 石垣司、サービス工学と時系列・多変量データ解析、非線形時系列に対する現象数理学の発展シンポジウム, 11 月 19~20 日, 東京, 2009 (依頼講演)

#### 6. 研究組織

(1)研究代表者

石垣 司 (ISHIGAKI TSUKASA)  
独立行政法人産業技術総合研究所・サービス  
工学研究センター・産総研特別研究員  
研究者番号：20469597

(2)研究分担者

なし

(3)連携研究者

なし