

投稿論文

通信販売事業における後払い決済の債権評価方法

筑波大学大学院システム情報工学研究科 東 弘樹

筑波大学大学院ビジネス科学研究科 津田 和彦

【要旨】

通信販売事業において、後払い決済を導入することは顧客満足度を高めると共に、売上向上にも貢献する。けれども、一部の顧客による支払遅延や未払いなどの問題が生じている。そのため、注文をした顧客が代金を支払うか否かを判断することは、注文を受けるか否かを判断の重要な要因となる。本論文では、通信販売業における後払い決済の債権評価方法を提案すると共に、評価用システムの構成を提案する。

後払い決済の債権評価方法は、主に2つの機能より構成する。1つは、注文情報チェックである。注文商品や代金総額、電話番号、注文主と送付先の関係などの支払状況から自学習するアルゴリズムを提案する。もう1つは、名寄せ機能である。注文方法に対して自然言語処理技術を用いることで、名寄せを行い、異なる記載方法でも同一人物の注文か否かを判定する。

これらで注文情報を評価することで、高い精度で未払いとなる注文を判別することが可能となった。

キーワード 後払い決済、評価システム、自然言語処理、統計処理、自動学習

【執筆者の担当箇所】

研究全般を主執筆者の東弘樹、分析の指導と助手および研究全般を東弘樹氏の指導教員である共著者の津田和彦が行った。

1. はじめに

日本における通信販売業界は、Webやモバイルなど申込手段の拡充と共に、「失われた20年」と呼ばれるバブル崩壊後の長期不況時代においても、その売り上げは上昇の一途を辿っている。公益社団法人日本通信販売協会（2015）の調査によると、2014年度（2014年4月－2015年3月）の通信販売市場の売上高は6兆1,560億円となり、10年前の2014年度の2倍を超える高い成長率となっている。

平成24年4月に発表された経済産業省（2012）第1回産業構造審議会流通部会審議用参考資料「我が国流通業の現状と取組・課題について」の中では日本における小売業の主要業態として、百貨店、総合スーパー、コンビニ、ドラッグストア、ホームセンター、そして通信販売の6業態が定義されている。

これら6業態の中で、通信販売を除く5業態は基本的に対面販売である。しかし、通信販売のみが顧客とは一切会わないという異なる事業形態となっている。それゆえ、通信販売業者の中には粗悪な商品を送付するなど悪質な業者が存在することは事実である。このことから通信販売業者自体が、顧客から十分な信頼を得られていないという課題がある。

この課題への対処方法として、いわゆる「後払い決済」を多くの通信販売業者は採用している。この決済方法は、顧客から見ると商品が届いてから支払を行うため、商品を確認した後に代金を支払えばよい。そのため、届いた商品が粗悪であったなどのトラブルに巻き込まれることが回避できるため、顧客からの信頼を得られる。

一方、後払い決済を取り入れた通信販売業者から見ると、顧客からの信頼感を得られるため、売上が向上するという利点がある。しかしながら、代金の支払いは顧客の良心に依存しているため、支払時期の遅延や未払いなどの問題が生じる。

そこで、本研究では不正購買顧客に対する通信販売業者の対応の実情調査を行い、課題を明確にすると共にその対策方法を提案する。具体的には、顧客からの注文情報を受け取った時点で分析を実行し、注文情報を評価する手法を提案する。さらには、提案手法を実行するシステムの構築方法を提案すると共に、実際に実験システムを構築し、提案手法の評価を実施した。その結果、注文情報を受け取ってから1秒以内に、その注文の支払可能性を評価することができた。

2. 後払い決済の債権管理

本章では、通信販売業者の債権管理に関する先行研究について概観すると共に、実際の通信販売業者による後払い決済の債権管理方法についての調査結果を示す。

2.1 通信販売における債権管理

通信販売には、様々な決済方法が存在している。主な支払方法と特徴を図表1に示す。

このように、通信販売における決済は多種多様である。経済産業省（2012）の第5回 産業構造審議会 流通部会 審議用参考資料によると、複数回答可能な調査において、2005年と2010年の決済方法の比較を行っている。

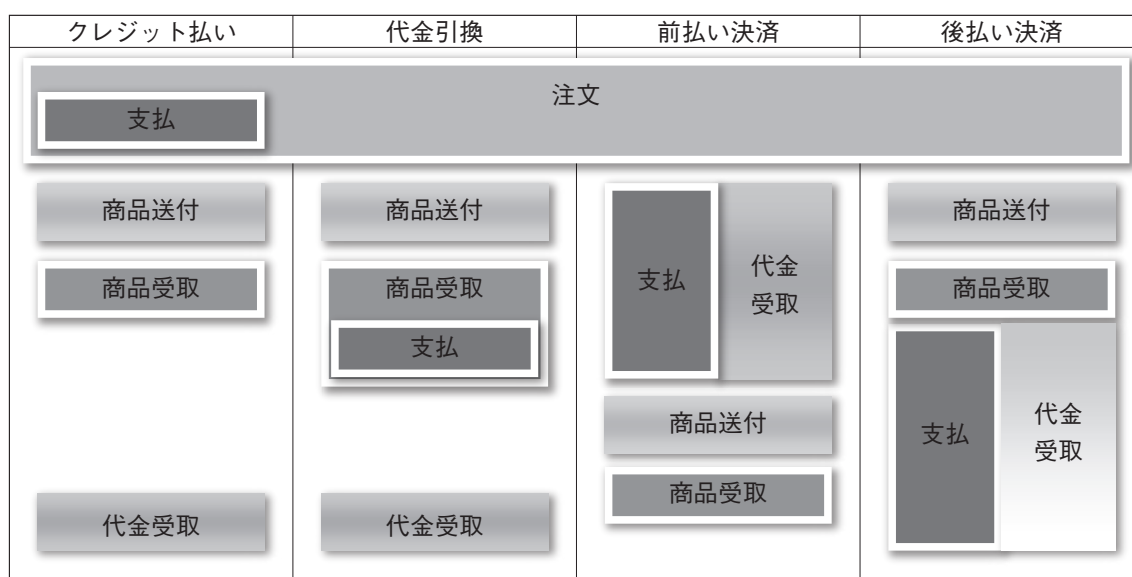
その結果、クレジットカード決済が増加、銀行振込・郵便振替が減少しており、代金引換、コンビニ決済は微増という状況を報告している。この中で、コンビニ決済、銀行振込・郵便振替では、前払い決済と後払い決済が存在するが、その大半は後払い決済である。

これら決済方法ごとに入金のタイミングなどが異なるため、通信販売業者にとっては、債権の管理方法が多種多様かつ複雑になるという課題が生じる。これらの決済方法を、通信販売業者から見た商品の発送や入金などのタイミングなどから再整理したものを図表2に示す。図表中、時間軸は上から下へ流れている。図表2からも明らかなように、後払い決済のみ顧客から見ると商品を受け取った後に支払を行うという形態になる。

図表1 通信販売における決済方法

決済方法	特徴
クレジットカード決済	注文後、カード会社が代金を立替払いした後に通信販売業者が商品を送付
代金引換	注文後、配送されてきた商品と引き換えに宅配業者に代金を支払う
コンビニ決済	前払い決済 注文後、通信販売業者から送られてきた振込用紙にてコンビニ店頭で支払い、通信販売業者が入金確認後商品を送付。
	後払い決済 商品と同梱して振込用紙が送付され、その振込用紙にて支払い
銀行振込・郵便振替	前払い決済 注文後銀行・郵便局で支払い、通信販売業者が入金確認後商品を送付。
	後払い決済 商品と同梱して振込用紙が送付され、その振込用紙にて支払い

図表2 通信販売における決済方法毎の購買イベントの発生タイミング



売上および売上債権管理に関する研究は、会計学に属する。通常、会計学における売上とは、商品もしくはサービスを提供した時点で計上される。その意味では、後払い決済のみはこれまでの会計学で扱われている事例に当てはまらない。会計学にも売掛・買掛という代金の支払タイミングが商品もしくはサービスの提供を終えた後という考え方はある。けれども、掛金は商品もしくはサービスを提供した時点で支払が約束されたことを前提としている。それゆえ、帳簿へは商品もしくはサービスを提供した時点で計上される。

通信販売業界についての関連研究は、受注前の活動に関しては、Simester（2006）による通信販売カタログの送付先の最適化モデル、君島（2010）による通信販売における広告投入と広告効果タイムラグを考慮した広告効果モデル、Conlin（2007）による通販受注と気温との関係进行分析、松田ら（2001）による通信販売業界における受注予測モデルなど、多くの研究が存在する。しかし、後払い決済の債権管理に関しては、東ら（2013）による通信販売における顧客の名寄せ知識の構築によるもののみで、会計学において通信販売における債権管理に関する先行研究は存在しない。

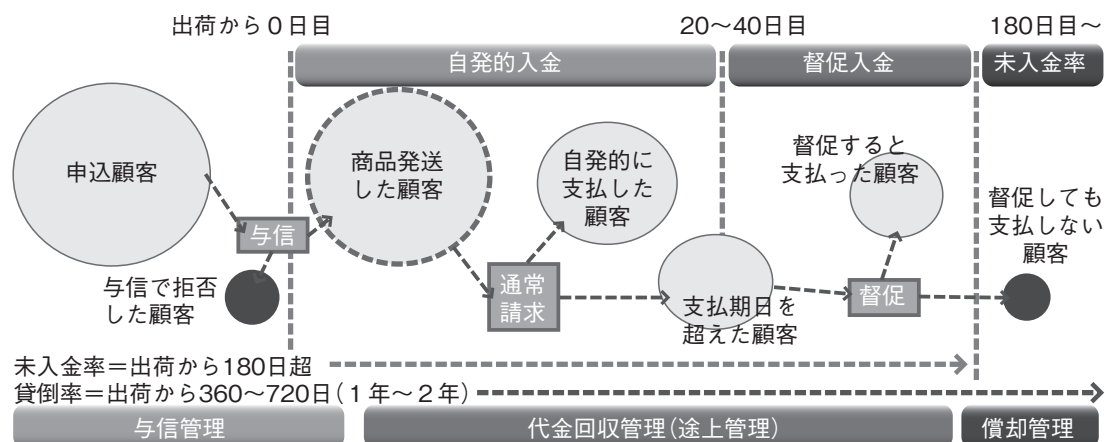
2.2 通信販売業による後払い決済の債権管理

通信販売業者における後払い決済の債権管理フローを、一般化したものを図表3に示す。後払い決済の場合、出荷時に債権が発生するが、その後、多くの企業は支払期日として2週間程度を設定している。何故なら、商品が到着や入金確認までにタイムラグが生じるからである。また、支払期日までに支払わない顧客もいるため、債権管理は年単位と長期間になる。

また、多くの通信販売業者では、支払期日までに支払わない顧客に対して、商品代金の支払いを則する督促を3回から6回程度行っている。そのため、常に債権の確認作業を行わねばならず、大きな負担となっている。

さらには、督促を繰り返しても商品代金支払われない債権は、弁護士法人やサービスなどに債権の回収を依頼すると共に、その注文顧客をブラックリストと呼ばれる今後発注を受け付けない顧客リストに追加する。それでも支払われない債権は、1年もしくは2年後に「貸倒」として会計上の処理を行う。これで、ようやく商品代金に関する債権の管理が完了する。

図表3 通信販売業者における後払い決済の債権管理フロー



図表4に商品の出荷数と支払状況についての事例を示す。図表中左側は、出荷数1000に対する入金率の推移であり、図表中右側は、左側の入金率モデルに最終的に未払となる出荷10を加えた場合の入金率の推移である。表中、入金率とは、各々の段階における入金率を示している。

図表4の左右で、支払期限までに入金される通常入金率には大差はない。けれども、督促回数を進めていくにつれ、その入金率の差異は顕著に表れる。このように、通信販売業者は面倒な後払い決済の債権管理を実施しているが、入金率が高いか低いかを判別するためには、多くの日数が経過しなければ判別できない状況となる。

通信販売業者にとって、後払い決済の債権管理業務はコストとなる。督促には、メールや電話、振込用紙の再発行・督促状の送付など様々な手段が用いられている。たとえば、図表4に示す事例において、督促を督促状の送付と仮定する。督促状の送付には概ね100円程度の直接経費が必要となる。その場合、図表4の左右を比較すると、左側は57通、右側は87通の督促状を送付したことになる。その差は30通なので、計3,000円余分な経費がかかったことになる。

通信販売業者において、後払い決済の債権管理における最終目的が顧客の支払であることは言うまでもない。けれども、上記のような煩雑な債権管理を行っても、早期に入金率が減少していることにも気付くことはできない。さらには、回収コストが上昇することも抑えることはできない。入金率を向上させるためには、注文を受けた時点での注文情報の評価が重要となることは明らかである。

図表4 未払となる出荷数の差による入金率の差異

	入金率	未払数	支払累計	未払数	入金率
出荷数	1000			1010	
通常入金	90%	100	900	110	89%
1回目督促	70%	30	970	40	64%

2回目督促	50%	15	985	25	38%
3回目督促	20%	12	988	22	12%
弁護士など	50%	6	994	16	27%
最終未払率	99.4%			98.4%	

3. 後払い決済の債権評価方法

前章で示したように、通信販売業者において後払い決済の注文を受ける際に、未払いとなる顧客からの注文を排除できれば、未回収による損失だけでなく、代金回収に伴う経費の削減など、その効果は非常に多大になる。本研究では、通信販売業者協力のもと、注文情報の調査を実施した。本章では、支払われる注文情報と未払いとなる注文情報との差異を明らかにすると共に、提案する注文情報の評価方法について説明する。

3.1 未払い注文の特徴

本節では、後払い決済にて注文された注文情報を分析した結果、明らかになった支払われる注文情報と未払いとなる注文情報との差異について記す。

注文情報の中には、主に図表5に示す情報が含まれている。送付先情報が複数存在する注文もあるが、多くの注文は1つであった。また商品情報は、⑨～⑫が複数存在するものは、多く存在している。

図表5 注文情報の項目

注文者情報	
①住所	
②氏名	
③電話番号	
④メールアドレス	
⑤その他 年齢・性別など	
送付先情報	
⑥住所	
⑦氏名	
⑧電話番号	
商品情報	
⑨商品名	
⑩数量	
⑪単品価格	
⑫合計価格	

以下、各項目の特徴について記す。

① ⑥住所：

通信販売業者ごとで、都道府県・市区町村・町名・番地・建物名と分割しているデータと一括のデータが存在するが、都道府県・市区町村・町名までの記載では地域別の入金率の差異は確認できたが、記載方法と入金率の関連は見出せなかった。一方、番地・建物名

では、記載方法により入金率に差異がみられた。たとえば、日本における集合団地の部屋番号の記載においては「103号室」となるのが一般であるが、「1F3号室」など明らかに違和感のある記載などは、入金率が低いという傾向がみられた。

その他、未払いとなった注文情報を列挙すると、明らかに同一場所に配達されるであろうが、住所記載方法が異なるものが存在した。この点に関しては次節にて詳細と解決方法を記載する。

② ⑦氏名：

日本にはおよそ30万の姓と100万弱の名が存在する。各々の姓の数と入金率の関係について分析を行った。その結果、「金」や「王」など外国人の姓の注文に入金率が低くなる傾向がみられたが、それ以外に特徴は見出せなかった。

③ ⑧電話番号：

電話番号においては、桁数間違いなど明らかに存在しない番号を記載した注文情報の入金率は低いという傾向が導かれた。間違いのないデータに関しては、自宅電話、携帯電話、IP電話の順で入金率が高いことが分かった。

④ メールアドレス：

メールアドレスにおいては、有料の特定プロバイダアドレスでは入金率が高く、YahooやGoogleなどのフリーメールは入金率が低い傾向がみられた。また、大手携帯キャリアの発行するメールアドレスは入金率が高い傾向がみられた。けれども、その差異は微小であった。

⑤ その他 年齢・性別など：

この情報を有する注文情報が少なかったため、十分な分析はできなかったが、高齢者ほど入金率が高い傾向がみられた。

⑨⑩⑪商品名・量・単品価格：

商品名については、Yahooオークションサイトに掲載されている商品と比較を行った。その結果、Yahooオークション掲載数と入金率には、逆相関の関係があることが分かった。その他、化粧品やパソコン関連部品など買取業者が存在する商品や、米や醤油など生活必需品は入金率が低い傾向もみられた。単品価格については、1万円前後の商品の入金率が低いという傾向がみられた。

⑫合計金額：

合計金額と入金率については、上限価格の20%から70%～80%の価格には金額が上がれば入金率が下がる傾向が確認できた。上限価格の70%～80%を超える注文に関しては、入金率は高かった。この原因は、通信販売業者ごとに設定されている上限価格ギリギリの注文情報に対してはチェックを行っていることに起因していると考えられる。

3.2 未払い注文の特徴による債権評価方法の優位性

東ら（2013）の主張によると、未払い者には、初めから支払意思のない悪質な顧客が存

在する。その特徴として、氏名や住所の表記の揺れを利用し、同一人物と特定されにくい手法をとる。マンション名の表記をしなかったり、表記を揺らしたりする特徴的な注文情報の支払済・未払いの記述率を検証した。支払済顧客と未払い顧客の割合を計算し、それぞれの割合を算出した。その一例を図表6に示す。

図表6 支払済顧客と未払い顧客の差異

	金額（百万円）			記述率		倍率
	Total	支払済	未払い	支払済	未払い	
平均値	1,771.68	1,687.66	84.015			
マンション名	2.102	1.934	0.168	0.1146%	0.2000%	1.74
名前ひらがな	0.446	0.412	0.034	0.0244%	0.0405%	1.66
様方	11.398	10.635	0.763	0.6302%	0.9082%	1.44
住所のひらがな・カナ・英字	2.991	2.804	0.187	0.1661%	0.2226%	1.34
字、ノ等の使い方	62.921	59.821	3.1	3.5446%	3.6898%	1.04

その結果、「マンション名」の記載が無い場合は、未払い率が1.74倍になるなど、支払済顧客と未払い顧客では記載方法の差異があることが判明した。また、悪意を持つ注文にも流行のようなものが存在し、抽出した名寄せを困難とする記載方法は、特定パターンが特定の時期に集中している場合が多いこともわかった。

3.3 後払い決済の債権評価方針

以上、注文情報に記載される各項目の特徴について解説した。その結果、注文情報から入金率の高低を評価するには、電話番号、合計金額、商品名、住所の4項目をチェックすれば効果的であることが分かった。

電話番号については、上記③⑧で示す事項のみのチェックで十分である。

合計金額に関しては、1注文であれば、各通信販売業者が設定する上限金額、売れ筋商品の価格や注文価格のモード値の間に設定すればよい。けれども、詐欺的な注文には、上限金額を回避するため、異なる人物からの注文を装った複数注文を行うという手口が存在する。これについては、3.4.1項に示す住所名寄せ手法で解説する。

商品名に関しては、前節で示したように都度Yahooのオークションサイトなどを閲覧する手法では、手間もかかると共にシステム化するコストは多大となる。さらには、住所記述においても、前節①⑥に示すように、明らかに誤魔化すことを目的とした記載方法はチェックする必要がある。

そこで、本研究では、商品名記述欄および住所記述欄の記載と支払情報の関連性を、人工知能における学習方法を応用したチェックポイント学習法を提案する。チェックポイント学習法については、3.5節で解説する。

3.4 住所名寄せ手法

通常、通信販売業者では、未払いとなった顧客の情報はブラックリストと呼ばれるリストを作成し、ブラックリストに掲載された顧客には商品を販売しないようにしている。また、多くの通信販売業者では、連続して複数の注文をする顧客もチェックしており、1人の顧客が通信販売業者の設定する上限を超えないように管理している。

ところが悪意を持つ未払い顧客は、ブラックリストに掲載されても、通信販売業者が同一人物からの注文と判断できないよう、注文情報の氏名や住所、電話番号などの記載を揺らして記載するなどの手口を使って、再度注文を行っている。また、通信販売業者が設定する上限を超える注文を行うためにも同様な手口が用いられている。

悪意を持つ未払い顧客の目的は商品の取り込みである。悪意を持つ未払い顧客は「荷物は受け取りたいが、本人特定されたくない」が本音である。通信販売で商品を購入する場合において、よほどの事情がない限り電話がかかってくることはない。また近年は表札を出さない家も多いため、氏名は重要な配送情報にはなり得ない。よって、電話番号や氏名はどのようにでも記載できることとなる。荷物を受け取るのに最も重要な情報は住所である。住所記述は、最終的には配送員が理解できる程度を超えた記載をすると荷物は受け取れない。それゆえ、その記載方法を揺らすことにも自ずと限界がある。

そこで、本研究では、住所記載に対する住所名寄せを行う手法を採用した。住所名寄せの手順は、3.4.1住所表記の揺れ除去、3.4.2表記の異なり度合いの評価の2つの過程で実施する。

3. 4. 1 住所表記の揺れ除去

まず、住所表記は通信販売業者ごとで、都道府県・市区町村・町名・番地・建物名と分割しているデータと一括のデータが存在する。分割しているデータでは、町名内に番地を記載するなど、あえて記述欄を間違えた記載が散見された。そこで、都道府県・市区町村・町名・番地・建物名と分割しているデータは、一つのデータとして連結する作業を実施した。

次に、日本語における異表記について統合する。具体的には、マンション名などの部分が主な対象となる。たとえば「ベレッジハイツ」などの「ベ」は「ヴェ」という異表記がある。その他、英文字の全角・半角・大文字・小文字、数字の全角・半角・漢字表記、など文字単位での揺れ除去を行った。

最後に、住所記述特有の表記揺れの除去を行う。福島県福島市のように、日本には都道府県名と市区町村名が同一の地域が多々存在する。この場合都道府県名を割愛する記載が多くみられた。そこで、都道府県のない住所記載については、都道府県を付与した。現在市区町村数は、1,718であり、都道府県との紐付は大きな手間にならなかった。また、町名に続く丁目・番・号・集合住宅の部屋番号などは、ハイフオンでつながれた数字列に略される。たとえば、正しい表現では「住吉町2丁目3番4号ベレッジハイツ105号室」は「住吉町2-3-4-105」などである。これらは、比較用文字列としては、丁目・番・号・

集合住宅の部屋番号などの番地部分をハイフオンでつながれた数字列に置き換えたものを生成して対処した。

図表7 住所記述の揺れ除去手順

福島市住吉町2丁目3番4号ヴェレージハイツ105号室	
福島県福島市住吉町2丁目3番4号ヴェレージハイツ105号室	都道府県名の付与
福島県福島市住吉町2丁目3番4号ベレージハイツ105号室	カタカナ表記統一
福島県福島市住吉町2-3-4-105	番地のデジタル化

3.4.2 表記の異なり度合いの評価

表記の異なり具合の評価に関しては、基本アルゴリズムとしてリーベンシュタイン距離を採用した。リーベンシュタイン距離とは、2つの文字列がどの程度異なっているかを示す距離算出方法の一つである。たとえば、一方の文字列に対して1文字の挿入や削除、置換などの操作により、もう一方の文字列と同じ文字列にするのに必要な操作の最小回数を2つの文字列の距離として定義する手法である。

この手法には、操作に重みを持たせることや、操作する対象を文字ではなく形態素とするなど様々な拡張が考えられる。たとえば、置換処理は削除と挿入の2つの操作とみれば、削除と挿入に対して置換は重み2とするなどである。

本研究では、リーベンシュタイン距離の算出方法として、第1段階の操作対象を形態素とした。また、重みはシソーラス辞書における意味的な類似度とした（山口翼、2003）。さらに第2段階として、文字を操作対象として距離を算出した。第2段階においては3.4.1で示すカタカナや英字、数字の統一処理も距離算出の要素とした。

本研究で提案した表記の異なり度合いの評価方法を用いて、ダミー住所に対して表記の異なり度合いを評価した結果を図表8に示す。図表8は、2行で1組になっている。図表8中、第1列が“-”となっているのが比較対象の住所記述で、その直下の行で第1列に記載された数字が、直上の比較対象住所との距離を類似度として示している。

図表8 住所名寄せ手法による表記の異なり度合いの評価事例（ダミー住所）

—	杉村泰介	香川県	善通寺市	仙遊町	3-1	
85	長瀬結花	香川県	善通寺市	仙遊町	3-1-5	
—	都築瑠衣	大阪府	堺市南区	別所	4-11-11	別所パール3004
90	中島御喜家	大阪府	堺市南区	別所	4-11-1	別所パーク304
—	古谷弥太郎	滋賀県	長浜市	新庄馬場町	4-3-1	パレス新庄馬場町413
93	大河原直治	滋賀県	長浜市	新庄馬場町	4-3-1	カレス新庄馬場4F3
—	山口咲良	高知県	香南市	香我美町上分	3-1-3	香我美町上分ランド302
96	石山満雄	高知県	香南市	香我美町上分	3-1-3	香我美町上分ランド3F2
—	高梨嘉子	鹿児島県	南九州市	川辺町永田	3-11-20	永田町ステーション32
96	濱田善之	鹿児島県	南九州市	川辺町永田	3-11-20	川辺町永田ステーション203

—	杉村啓之	岐阜県	大垣市	小泉町	1-8-15	パークスナイン
100	大村芳雄	岐阜県	大垣市	小泉町	1-8-15	パークナイン
—	西川文香	鳥取県	八頭郡智頭町	大背	3-14	ビレッジ大背209
100	今西由真	鳥取県	八頭郡智頭町	大背	3-14	ヴィレッジ大背209
—	笹川藤子	宮崎県	宮崎市	糸原	1-12-13	リバーサイド糸原109
100	亀井穂乃香	宮崎県	宮崎市	糸原	1-12-13	リバーサイドitohara109

3.5 チェックポイント学習法

3.5.1 パトリシアトライ

獅々堀正幹ら（1997）によると、大規模なテキストを扱う情報検索や自然言語処理では、文字単位で高速に検索できるデータ構造として、トライ（Trie）が索引や辞書を構成するために利用されている。トライはデータを構成する文字を分岐条件とした木構造であり、特に0,1のみを分岐条件とするトライを2進木トライ（Binary trie）という。

パトリシアトライ（Patricia trie）は、分岐を持たない内部節を取り除いた2進木トライであり、取り除いた節の数を各内部節に保存しておく。検索時には分岐のある内部節のみを参照するため、通常の2進木トライよりも高速に探索を進めることができる。また、節の数が大幅に削減されるため、よりコンパクトに索引や辞書を実現できる。

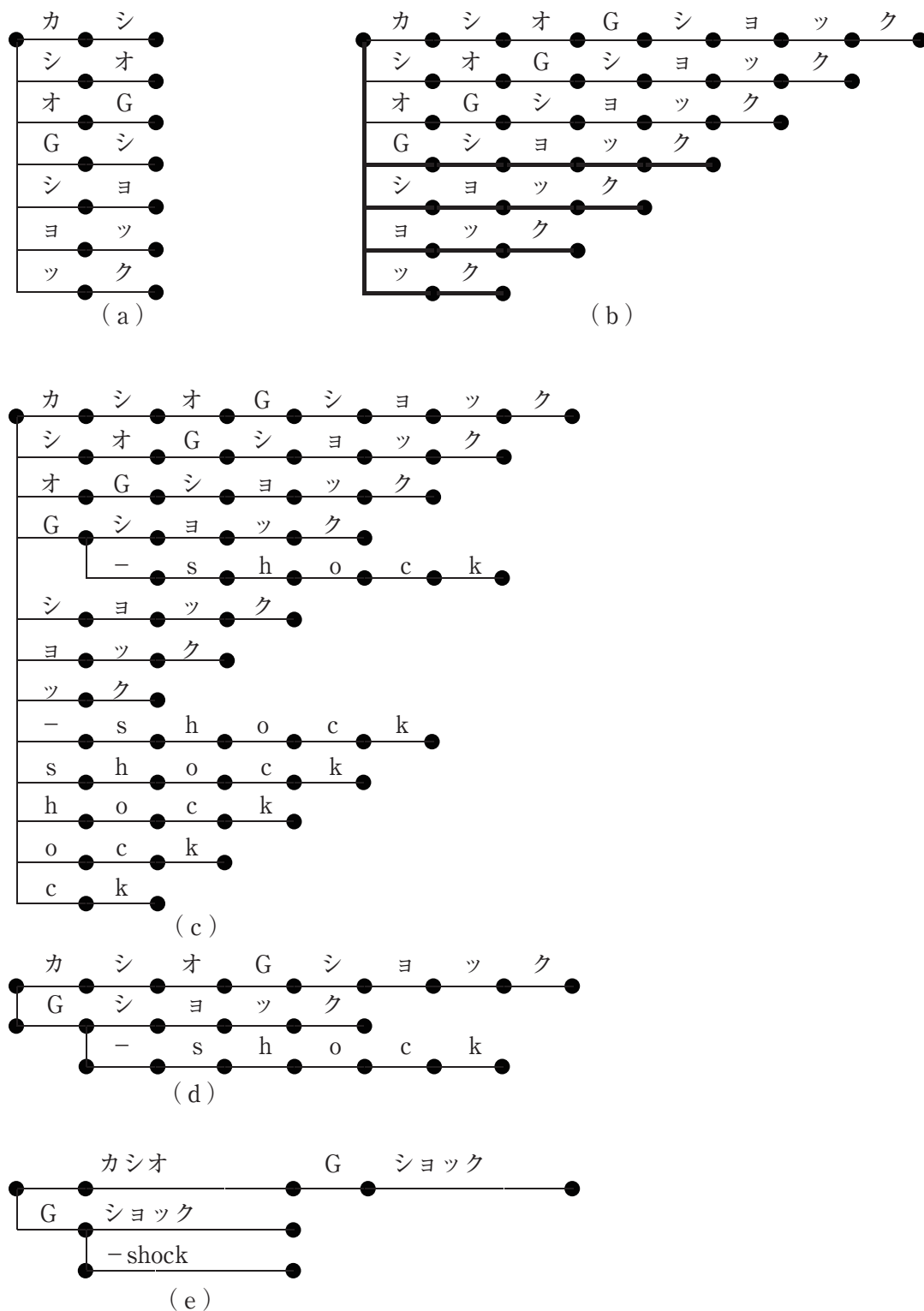
3.5.2 パトリシアトライを応用したチェックポイント学習法

通信販売業者が取り扱っている商品は、多種多様である。また新商品の発売や廃品になるものなど、商品の入れ替わりも激しい。また、住所などの記載も新たな手口が次から次へと出てくるといった状況がある。

そこで、商品情報については、商品名の文字列についてパトリシアトライを用いて部分文字列を抽出する。そのうえで、抽出した部分文字列が含まれる注文の支払状況の差異を集計して、その部分文字列の評価ポイントを算出する。そして、注文の商品情報の中にその部分文字列が含まれるか否かを照合し、含まれていた場合に算出した評価ポイントにより、商品情報を評価する手法を用いる。

この手法の利点は、商品名などを意識せずに機械的に処理できる点である。それゆえ、汎用性も高い手法である。欠点は、計算量が非常に大きくなることである。けれども、現状のコンピュータの計算能力から考えると、大きな負荷にはならない。

図表9 パトリシアトライの構築手順



パトリシアトライの構成方法について {“カシオ G ショック”、“G ショック”、“G — shock”} の3つのキーワードから構成する手順を説明する。

まず通常のトライを構築する。トライとは、図表9にて●で示すノードと、線で示すエッジから構成され、エッジに文字ラベルを付与したものである。

“カシオGショック”という文字列に対して、全2文字接続を取り、トライを構築する。この状態が図表9(a)となる。その後、全3文字接続を取り(a)に追加、以降、全4文字接続、全5文字接続と順に接続文字列を長くしていき(a)に追加していくと、最終的に図表9(b)の状態となる。次に、図表9(b)上に“Gショック”を先と同じ手順で追加する。“Gショック”自体が“カシオGショック”の部分文字列となっているため、図表9(b)の太線部分となり完全に重なった形になる。続いて、図表9(b)上に“G-shock”を先と同じ手順で追加する。図表9(c)の上4行目のエッジには既に‘G’があるので、‘G’の後に枝分かれして“- s h o c k”の枝が追加される。それ以外は“- s h o c k”に対する部分文字列の枝が追加され、最終的に図表9(c)に示すトライが完成する。

図表9(c)に示すトライが完成すると、まず、始点から終点までの枝の文字列が、他の枝の部分文字列となっている枝を全て削除し、図表9(d)に示すトライにする。その後、短い枝から順に分岐のない枝の文字を結合し文字列にしていくと共に、短い枝にある文字列と同じ文字列が存在する長い枝は、短い枝にある文字列のみ分割した文字列の集合体に区切り、図表9(e)に示すパトリシアトライを完成させる。

このパトリシアトライより、キーワードとしては{“カシオ”、“G”、“ショック”、“-shock”、“カシオG”、“Gショック”、“カシオGショック”、“G-shock”}の8つが抽出される。そして、各々のキーワードが含まれる全注文情報から、その入金率を算出する。

住所欄についても明らかに操作していると思われる記載が見受けられる。たとえばマンションなど集合住宅の部屋番号について、本来は「101号室」と記載すべきところを「1F 1号室」とか「1階右」とかと記載するなどである。これらについても、住所の町名以降の記載部分に対して、商品情報の際と同じパトリシアトライを、1文字を部分文字列の最小単位として構築し、当該注文の支払率よりチェックポイントを抽出する手段を適用した。

メールアドレスに対しては、“@”と“.”で区切りがあるため、区切り間の文字列を最小単位として構築し、当該注文の入金率よりチェックポイントを抽出する手段を適用した。

氏名に関しては文字数が概ね4～6文字と短いことより、電話番号に関しては数字列のみでパターンが少ないことより、パトリシアトライの効果は低いため適用外とした。

4. 後払い決済の債権評価システム

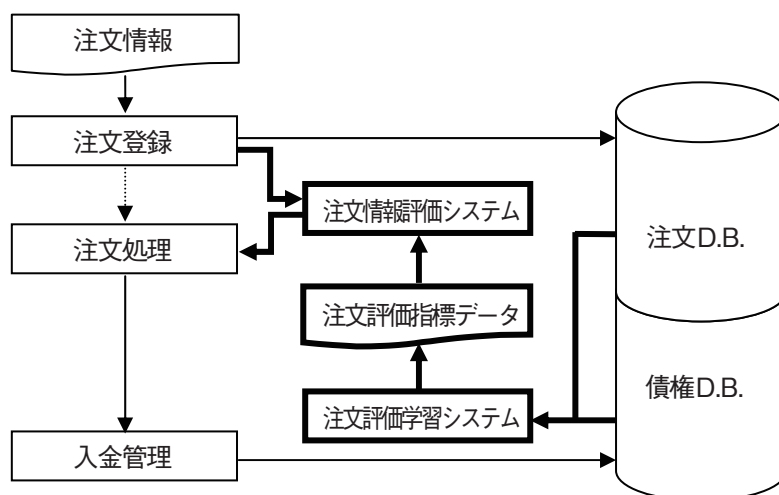
本章では、3章に示した手法を実現するシステム構成を提案すると共に、その評価結果を記す。実際の通信販売業者の協力を得て、3章に示す手法を用い実際の注文情報を評価するシステムを構築すると共に、検証を実施した。その詳細について述べる。

4. 1 システム構成

本論文で提案する後払い決済の債権評価システムを、通信販売業者の注文&債権管理を行うシステムへ組み込んだ場合のシステム構成を図表10に示す。本論文次節に示す評価も、図表10に示す構成で通信販売業者の注文&債権管理システムへ後払い決済の債権評価システムを組み込んで実施した。

図表10に示すシステム構成の詳細を説明する。図表中、左部の上部から下部への流れは、注文を受けると、注文情報をデータベースへ登録した後、注文商品の発送、債権情報の生成など注文処理を行う。注文した顧客が注文商品を受け取ると、入金管理を実施する。入金管理では、指定期日までに支払いがなかった場合には督促を行う。多くの通信販売業者では、商品出荷から半年程度の期間に3回から6回程度の督促を行う。その後に、弁護士などへ債権回収の依頼を行い、1年から2年程度経過しても

図表10 後払い決済の債権評価システムを組み込んだ注文&債権管理システム構成



未入金となる債権は、特損処理をして帳簿から消し去っている。

本論文で提案する後払い決済の債権評価システムは、図表10中太字枠で示す部分である。後払い決済の債権評価システムは、大きく分けて注文情報学習システムと注文情報評価システムの2つのブロックで構成される。

図表に示すように、後払い決済の債権評価システムがアクセスするのは、どこの通信販売業者が有する注文データベースと債権データベースである。そして、注文登録の後、注文処理を行うまでの間に注文情報評価システムを介するようにフローを変更する。以上で、通信販売業者が有する既存のシステムへ組み込むことができるので、本論文で提案する後払い決済の債権評価システムは、適用性が高いシステムと言える。

注文評価学習システムでは、注文データベースと債権データベースを閲覧し、注文情報を評価するための情報を抽出する。具体的には、未払い注文リストと、3.5節で示すチェックポイント学習法で示した手法を用いて抽出したチェックポイントである。この2つの

情報を注文評価指標データとして出力する。

注文情報評価システムは、注文を受ける都度、評価を行う。具体的には、名寄せ処理を用いて評価対象注文と注文評価指標データにある未払い注文リストとの照合、チェックポイントとの照合を行い、検出された事項から、注文の評価を数値化して出力する。

4. 2 評価方法

まず、実験環境について記す。

実験は、通信販売業者にご協力いただき、2015年5月中の特定期間¹⁾に受けた注文のうち、支払方法として後払い決済を指定した8,020注文を全て受注して頂き実施した。

4. 2. 1 未払いリストとの照合

未払い注文リストは、債権データベースから過去5年間分の未払い注文を抽出した。未払いリストとの住所照合結果 est_{add} は、式1にて数値化し評価する。

$$est_{add} = \sum k_{add} * sml \quad (\text{式1})$$

k_{add} は、未払い注文の係数である。未払い注文の係数は、前の注文が、前日注文、1ヵ月以内の注文、1ヶ月以上以前の注文の3つに分類した。前日注文は、買い忘れによる買い足しもあるため、1件目の未払い係数は0に設定した。前日の2件目以降および1ヵ月以内の注文は、過去1年間の未払い中の注文の入金率から算出し、0.4に設定した。1ヶ月以上以前の注文は、係数3とした。

sml は3.4.1項に示した住所名寄せにて算出した比較対象住所との類似度の百分率とした。ここでは、距離が80以上を対象とすることとした。

数式中の Σ は、未払い注文が複数ある場合は、その累積とするための処理である。

名前や電話番号、メールアドレスに関しては、完全一致するか否かで照合するが、これは同一顧客からの確認のためのデータとし、評価を行うこととした。

また、後処理として未払い注文の累積が、通信販売業者の設定する後払い決済枠の金額を超えているかもチェックし、枠を超えると est_{add} は自動的に3となるように設定した。

4. 2. 2 チェックポイントとの照合

商品名、住所、メールアドレスのチェックポイントは、3.5節で示すパトリシアトライを用いた手法で抽出した。抽出結果は（チェックする文字列、評価点 k_{chk} ）のセットのリストとなる。評価点は、「該当注文の未払い率²⁾／全注文の未払い率」で算出した。

電話番号に関しては 自宅電話、携帯電話、IP電話の分類のみを行い、入金率の悪いIP電話のみ上記と同じ評価点の算出方法で算出した。

そのうえで、チェックポイントとの照合結果 est_{chk} は、式2にて数値化し評価する。

$$est_{chk} = \sqrt{n} * \Sigma k_{chk} \quad (\text{式 2})$$

k_{chk} は評価点である。 Σ はチェックポイントと照合できた項目が複数あった場合に、その累積とするための書類である。また、 \sqrt{n} は、チェックポイントと照合できた項目数の平方根である。複数のチェックポイントに照合された場合、注文情報を不正に操作した可能性があるため、 \sqrt{n} を乗算することとした。

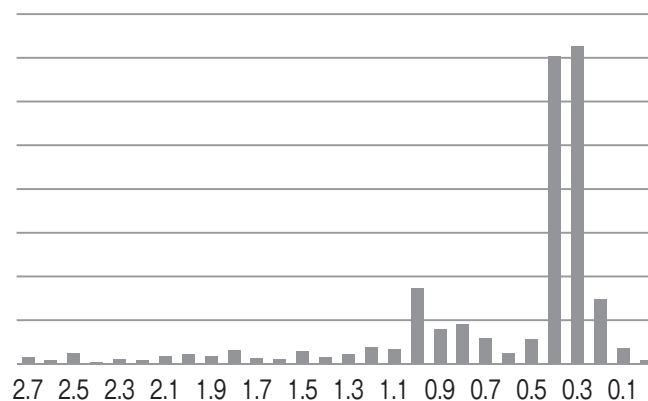
4. 3 評価結果

本実験による評価では、未払いリストとの照合において算出した評価結果 est_{add} と、チェックポイントとの照合において算出した評価結果 est_{chk} との和を、評価対象注文の評価値として採用した。その結果、評価値が最も良い注文は0となり、注文の評価値が大きくなるほど悪い注文となる。

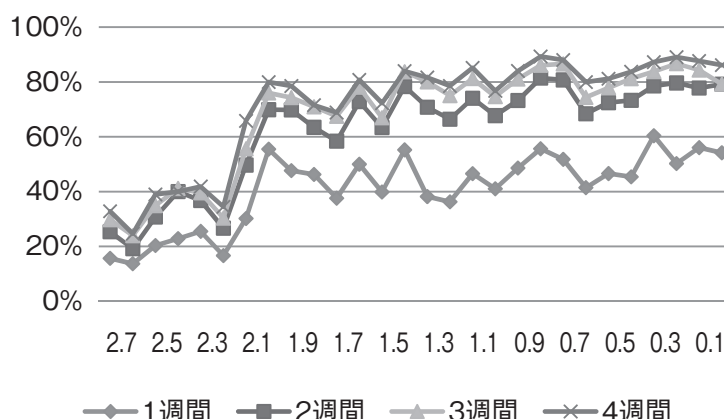
実験は、通信販売業者にご協力いただき、2015年5月中の特定期間に受けた注文のうち、支払方法として後払い決済を指定した8,020注文を全て受注して頂き実施した。

評価値は、最も良い評価の注文は0、最も評価の悪い注文は3.0となった。評価値の分布を図表11に示す。多くの注文の評価値は0.3前後となっている。評価値が0.5以下の注文で、全注文の70%を占めた。また、評価値が1.2以下で全注文の90%を占めた。

図表11 評価値の分布



図表12 評価値毎の4週間の入金率



次に、評価値ごとの入金率を見る。図表12には、商品出荷日を起点として、1週間目から4週間目の入金率を示す。4週間目としたのは、これ以降は督促を行うためである。よって図表12に示す入金率は、顧客自身が自発的に支払った割合となる。

4週間経過時点での全体の入金率は、約86.6%であった。この数値は、支払方法を後払い決済と指定した全注文を受注したことを考えると、悪くない数値である。図表12からも明らかなように、評価値が良いほど入金率が良いという結果となっている。このことから本論文で提案した注文情報の評価は正しいと言える。

多くの通信販売業者は後払い決済で申し込まれた注文の5%程度を、お断りもしくは支払方法を先払いへの変更をお願いしている。本実験の結果では、評価値が2を上回ると急激に入金率が低下している。評価値2を超える注文数は、全注文の約3.4%であった。そこで、評価値2を超える注文をお断りしたと仮定すると、入金率は約88.2%となる。

この数値に対して、数社の通信販売業者の債権管理担当者にヒアリングを行った。その結果、全く人手をかけずに督促前に得られる入金率としては十分に評価できる、との意見を得た。

5. おわりに

本論文では、支払方法として後払い決済を選択した注文情報を、受注時に評価する手法を提案すると共に、実際の通信販売業者の協力を得て実証実験を行った。その結果、十分に実用的な数値を得ることができた。

また、注文情報評価システムの処理時間も、1注文あたり1秒以内で評価結果を算出しており、十分実用的な処理時間と言える。

本実験と並行して、実験にご協力いただいた通信販売業者は、人手による注文情報のチェックも行っている。その結果、人手によるチェックが有効に働くのは、提案手法での評

価値が1.2以上となった約10%の注文であった。このことから、提案手法を導入することで、人手によるチェックが必要な注文情報数を大幅に削減できる可能性があることが示唆される。また、精度の高い注文情報評価システムは、売上に応じた調整が可能となる。例えば、「与信を緩くしてお断りする件数を減少させる」・「与信を厳しくしてお断りする件数を増加させる」事が可能となる。この時に一定の与信精度を保つために、上限額を上下させ事も合わせて検討する必要がある。

提案手法で評価値2という入金率が悪いと判定した注文情報のうち、人手によるチェックでは入金率が良いと判定した注文情報を調査した。提案手法の評価の根拠は商品情報であり、人手によるチェックでの評価の根拠はお得意さまであった。このことからお得意様情報を加味すると、さらなる精度向上が望めるであろう。

今後は、お得意様情報を加味すると共に、注文情報チェックの担当者からのヒアリングを積み重ね、さらなる精度向上に取り組む所存である。

【謝辞】

本研究は、平成26年度ダイレクトマーケティング学会研究プロジェクト（自主研究：通信販売事業における不正購買対策モデルの研究）にて実施した研究である。

【参考文献】

- ・ 青江順一；“トライとその応用”，キー検索技法—情報処理学会論文誌，34，2，pp. 244 - 251, 1993.
- ・ 東弘樹,津田和彦“通信販売における顧客の名寄せ知識の構築”，2013年度人工知能学会全国大会, 1F4-2, Jun.2013.
- ・ 東弘樹,高橋雅和,津田和彦“通信販売業界における督促効果測定”，電気学会第61回情報システム研究会, IS-14-054, Nov. 2014.
- ・ 東弘樹,高橋雅和,津田和彦“通信販売業における受注経路の代金支払傾向の知識化に関する考察”，2015年度人工知能学会全国大会, 2G5-OS-25b-2, May.2015.
- ・ 東弘樹,高橋雅和,津田和彦“通通信販売業における効果的な督促を行うための顧客区分抽出方法”，2015年度電気学会第63回情報システム研究会, IS-15-023, July.2015.
- ・ 東弘樹，藤田昌克，津田和彦“通信販売業界における代金支払いの可能性を評価するアルゴリズムの研究”，2015年経営情報学会秋季全国研究発表大会, PR0032, Nov.2015
- ・ 北研二，津田和彦，獅々堀正幹：“情報検索アルゴリズム”，共立出版，2002.
- ・ 経済産業省「第1回 産業構造審議会 流通部会 審議用参考資料」，Apr.2012.
- ・ 経済産業省「第5回 産業構造審議会 流通部会 審議用参考資料」，Jul.2012.
- ・ 中村康正，望月久稔. 圧縮デジタル探索木における辞書情報更新の高速化手法. 情報処理学会：データベース，Vol. 47, SIG 13 (TOD 31)，pp. 16-27, 2006.
- ・ 公益社団法人 日本通信販売協会「2014年度（平成26年度）通信販売売上高について」，

Aug.2015.

- ・ 矢田 普, 大野将樹, 森田和宏, 泓田正雄, 吉成友子, 青江順一. 接頭辞ダブル配列における空間効率を低下させないキー削除法. 情報処理学会論文誌, Vol. 47, No. 6, pp. 1894-1902, 2006.
- ・ 山口翼編, 日本語大シソーラス—類語検索大辞典, 大修館書店, 2003
- ・ Daniel Jurafsky and James H.Martin: Speech and Language Processing, pp.74, Prentice Hall, 2009, ISBN 0-13-187321-0
- ・ Gusfield, Dan (1997) . Algorithms on strings, trees, and sequences. Cambridge University Press. ISBN 0-521-58519-8. MR1460730. Zbl 0934.68103.
- ・ Hiroki AZUMA, Masakazu TAKAHASHI and Kazuhiko TSUDA, “Building Knowledge of Customer Distinction for Safe Transaction in the Mail Order Industry”, ASIC2013, Dec. 2013.
- ・ Hiroki AZUMA, Masakazu TAKAHASHI and Kazuhiko TSUDA, “A Study on the Effect Measurements of Credit Reminder in the Mail-order Industry”, ASIC2014, Dec. 2014.
- ・ Hiroki AZUMA, Masakazu TAKAHASHI and Kazuhiko TSUDA, “Understanding Customer Characteristic of the Payment Trends in the Mail Order Industry”, ASIC2015, Oct. 2015.
- ・ J. Aoe, K. Morimoto, M. Shishibori, and K-H. Park. A trie compaction algorithm for a large set of keys. IEEE Trans. Knowledge and Data Engineering, Vol. 8, No. 3, pp. 476-491, 1996.
- ・ kimijima M.: A Study on Measuring Input-Output Process on Order-Getting Costs for Direct Marketing, Yokohama National University Departmental Bulletin Paper, 16(1), 21-39, 2010 (Japanese)
- ・ D . Morrison : “PATRICIA – Practical Algorithm To Retrieve Information Coded in Alphanumeric ”, Journal of the Association for Computing Machinery , 15 , 4 , pp. 514 – 534, 1968.
- ・ D. R. Morrison. Patricia practical algorithm to retrieve information coded in alphanumeric. Journal of the ACM, Vol. 15, pp. 514-534, 1968.
- ・ Simester, D. I., Sun, P. and Tsitsiklis J. N.: Dynamic Catalog Mailing Policies, Management Science, 52 (5) , 683-696, 2006.

[2016年 1 月12日受付、2016年 2 月26日受理]

【注】

- 1) ご協力いただいた通信販売業者のご意向により、実験期間の公開は不可