

平成23年 5月31日現在

研究種目：若手研究 (B)
研究期間：2007 ~ 2010
課題番号：19730275
研究課題名 (和文) 大規模会話データを用いた説得コミュニケーションプロセスの解明
研究課題名 (英文) Analysis Framework for Persuading Process and an Application to Debt-Collecting Conversation Logs

研究代表者

矢田 勝 俊 (YADA KATSUTOSHI)

関西大学・商学部・教授

研究者番号：00298811

研究成果の概要 (和文)：

本研究の目的は大規模で所与の構造を持たないコミュニケーションデータから説得プロセスに関する有用な特徴を明らかにする枠組み、分析手法を提案し、企業の債権回収にかかわる会話データの解析を通じて提案手法の有用性を検討することである。本研究は会話の流れを定量化し、解析する手法を開発することで、債権回収における支払い遅延顧客を説得するコミュニケーションプロセスを明らかにした。

研究成果の概要 (英文)：

The purpose of this research is to develop a framework and an analysis method to discover some features from the content and the persuading process in un-structured databases, and to show its application to the communication for the debt-collecting process to verify the usefulness of the proposed system. In this study, we were successful in clarifying the communication process in the overdue payment collection by using the data of a Japanese telecommunication company.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	700,000	0	700,000
2008年度	800,000	240,000	1,040,000
2009年度	900,000	270,000	1,170,000
年度			
年度			
総計	2,400,000	510,000	2,910,000

研究分野：経営情報論

科研費の分科・細目：経営学

キーワード：データマイニング・経営情報・マーケティング

1. 研究開始当初の背景

マーケティングや経営分野において、テキストマイニングのビジネス応用に関する既存研究が存在する。アンケートの自由回答（上田 2005; 豊田 2002; Araki 2005; Yada 2006）や有価証券報告書（喜田 2006）などのテキストデータを対象にして、特定単語の使用頻度の基礎的統計指標をもとに、有用な知見を発見しようとしている。従来、特定のデータ構造を持たない文章などの解析では、内容分析（Krippendorff 1980）が社会学において開発されており、これらの研究は計算機性能の向上をてこに内容分析の対象を広げ、そのデータ規模を拡大させている。しかしながら、これらの研究は解析手法や枠組みについて従来の内容分析に依拠しており、分析手法またはコミュニケーション研究としての理論的な貢献に乏しい。一方で、AI（人工知能）研究におけるテキストマイニング研究として、田中ら（2005; 2006）による紛争処理に関する文章解析研究がある。これらは調停に伴うコミュニケーションデータから、類似の紛争案件を検索し調停支援を行い、アドバイスを提供しようとする試みである。これらは支援のタイミングを扱っているが、コミュニケーションのプロセスそのものを解析する枠組みを明らかにしていない。本研究では、コミュニケーションプロセスをモデル化し、非構造のテキストデータを定量化する枠組みを提案する。

私はこれまで顧客の購買履歴データに代表されるような企業内の構造化されたデータを解析する手法や枠組みを研究してきた。特に販売データの解析では、AMA や AMS などの国際会議で研究成果をあげることができた。一方でこうした定量化しやすいデータの限界も明らかになっており、近年、それを補完する文章などの定性的（非構造）データの解析に研究範囲を拡大している途上である。本研究はそうした定量的解析と定性的解析の統合を目指した第一歩である。また本研究の基礎的な考察として、Yada(2006)はテキストマイニングと因子分析を統合し、説得に用いられる単語の利用頻度の特徴について明らかにした。しかし、これは既存研究と同様に単語の利用頻度をもとにした静的な解析であり、会話中の時系列データはその一部しか扱っておらず、コミュニケーションプロセスに着目した動的な特徴を抽出することは難しい。

2. 研究の目的

本研究の目的は、大規模で所与の構造を持たないコミュニケーションデータから説得プロセスに関する有用な特徴を明らかにする枠組み、分析手法を提案し、企業の債権回

収に関わる会話データの解析を通じて提案手法の有用性を検討することである。従来、会話データのように構造を持たないデータの定量的解析は単語の利用頻度といった単純な解析にとどまるものが多い。本研究は、単語の利用頻度といった静的な文章の特徴だけでなく、動的なコミュニケーションプロセスの特徴を明らかにする枠組み、分析手法を提案する。企業の商品購入代金の支払い遅延に関して顧客と交渉する大規模な説得コミュニケーションデータを対象とする。これには、企業と顧客との膨大な会話データが記録されており、これらから説得プロセスに関する有用な知見を発見することが企業にとって重要な意義を持つ。つまり大規模な支払い遅延顧客との会話データから、かれらを説得するためのコミュニケーションプロセスを明らかにする。本研究は、説得コミュニケーションプロセスのモデル化から、解析システムの開発、実データによる実証までを行う包括的なテキストマイニング研究である。

3. 研究の方法

(1) 文献研究と枠組みの提案

社会学における説得コミュニケーション研究、AIにおけるテキストマイニング研究などの文献を中心にレビューし、コミュニケーションプロセスのモデル化を試みる。具体的には説得コミュニケーションにおける既存研究の成果に基づき、説得プロセスをいくつかのステップに分類し、それらの特徴を定量化する枠組みを提案する。

(2) 解析手法の開発

上記のモデル化にそった、新しい文章解析のためのアルゴリズムとシステムの開発を行う。特にプロセスの特徴の定量化と大規模データに耐えうる計算速度を実現する。文章解析のアルゴリズムは、広島市立大学の砂山研究室と共同開発する。

(3) データ収集・蓄積とシステム環境の構築

本研究の対象となる債権回収に関する説得コミュニケーションの会話データを記録、蓄積する仕組みを構築する。関係企業との協力の確約は既に取れており、社内におけるデータ収集と同時にヒアリング調査を実施する。ただし、そのデータの一部は音声（会話）データとして記録されているため、人手をかけて電子化する。

(4) 画像データの収集と統合

本研究では会話データだけではなく、説得交渉の際の状況を録画した画像データも解析対象に含める。人のしぐさや距離など画像から抽出される特徴をコミュニケーションプロセスの特徴として利用する。

(5) 提案手法の評価、理論的考察

蓄積された企業データを用いて提案手法

の有用性の評価を行う。企業内の専門家の評価、シミュレーションによる精度などを検討し、理論的な考察を加える。

4. 研究成果

(1) 理論的貢献

本研究の目的は大規模で所与の構造を持たないコミュニケーションデータから説得プロセスに関する有用な特徴を明らかにする枠組み、分析手法を提案し、企業の債権回収にかかわる会話データの解析を通じて提案手法の有用性を検討することである。本研究は会話の流れを解析する手法として、説得プロセスをモデル化し、ラベル付けを行うことで文章を時系列で定量化することに成功した。説得プロセスとして、行動内容の説明、行動意思の付与、行動手順の相談、行動の最終確認を定義し、それぞれの段階に関連する単語集合にラベル付けする。それらの出現頻度を時系列で解析することで、有用な知見が発見できる。本研究では、回収率の高いオペレーターの説得プロセスの特徴をこれらの変化から抽出することで、債権回収における支払い遅延顧客との大規模な会話データから彼らを説得するためのコミュニケーションプロセスを明らかにした。

(2) 実験データから得られた具体的知見

実験結果から、それぞれのプロセスに関連づけられた単語の使い方が、オペレーターの回収率に大きな影響を与えることが明らかになった。表1は説得プロセスにラベル付けされた単語の出現頻度と債権回収率との相関関係を示している。

表1 説得プロセスに関連する単語と回収率との相関関係

説得プロセス	関連語との相関
行動内容の説明	-0.46
行動手段の確認	-0.81

まず行動内容の説明に関する単語群が回収率にネガティブな影響を与えていることがわかる。債権回収の場合、債権の状況などを説明しているものであり、こうした会話が中盤まで続いているオペレーターは回収率が低下する傾向が見られた。一方、行動手段の確認とは、債権回収における回収方法つまり支払い方法などを指しており、具体的な返済手続きを話し合っていることを意味している。回収率の高いオペレーターは債務者を自らの会話ペースに巻き込むために、会話の流れを変える言葉をタイミングよく使い、会話開始後2-3分以内に具体的な支払い手続きに関する話し合いを始めることが分かった。

このように説得プロセスをモデル化し、定量化することで、債権回収に関する重要な知

見をえることができた。発見された知見を要約すると、①早い段階で相手に同調の意思を伝え、②相手の置かれている状況を素早く理解する、そして③相手の理解の程度を確認しながら具体的に支払い方法の相談を進めることが重要である。これらの知見をもとに、適切なオペレーターの研修プログラムを開発することができた。

(3) 進捗評価

本研究で開発したテキスト解析技術はゲノム領域で開発された文字列解析技術を応用したもので、多次元の時系列データを解析するのに有用である。研究計画では、最終年度にこれまでの研究解析のとりまとめと開発した手法の他領域への応用可能性を検討する予定であったが、環境整備のデータ収集が企業の協力によって順調に進み、開発や分析を当初の計画よりも前倒しで行った結果、砂山・矢田(2007)、Ohsawa and Yada(2009)など、国内外の学術雑誌や書籍に研究論文を公刊、当初の予定以上の研究成果を達成することができた。

(4) 今後の展望

当初の研究計画で予定していた「説得コミュニケーションプロセス」のモデル化に関する研究目標を達成し、引き続き、文字列解析技術をベースに開発した多次元時系列解析手法の他領域への応用可能性を検討していた。この手法は当計画の対象となるテキストデータだけではなく、ストリームデータにも有用な手法であると考えられ、近年のマーケティング研究において、顧客動線を追跡したストリームデータの解析が大きな注目を集めており、早急に当該研究に着手する必要があると考えたからである。しかし、対象がテキストデータからストリームデータに変わる事、データ規模が大規模になる事、そして店舗実験を実施しなければならないことなどから、当初予定していた研究経費では不足するため、研究計画最終年度前年度応募として、H22年度基盤研究(A)に応募し、その交付内定を受けるに至った。これにより今後は、小売業における顧客の購買履歴データとセンサー技術を用いて定量的に測定した実際の店内移動経路に関する顧客データを統合する事によって、顧客の店内における移動行為と購買結果の因果関係を明らかにする消費者行動モデルを構築し、実証実験から得られる現実の顧客データによって、その有用性を検証する。

<参考文献>

上田隆徳編、講談社、『テキストマイニングによるマーケティング調査』, 2005, 244p.
豊田裕貴, “テキストデータの数値化 —分析者の視点からのテキストマイニング前処理における注意点—”, 人工知能学会誌, Vol.1.

17, No. 6, 2002, pp738-743.

T. Araki, S. Hamada, N. Matsumura, S. Niwase, Y. Ohsawa and K. Yada, "Chance Discovery from Consumer Research Using KeyGraph", Readings in Chance Discovery, Advanced Knowledge International, 2005, pp. 373-384.

K. Yada and N. Matsumura, "Knowledge Discovery from the Structure of Persuasive Communication", Proc. of 2006 IEEE Conference on System, Man and Cybernetics, IEEE Press, 2006, pp.1741-1746.

喜田昌樹, "アサヒの組織革新の認知的研究—有価証券報告書のテキストマイニング—", 組織科学, 第39巻, 第4号, 2006, pp79-92.

K. Krippendorff, Sage Publication, Content Analysis: An Introduction to Its Methodology, 1980, 413p.

田中 貴紘, 安村 禎明, 片上 大輔, 新田 克己, "オンライン調停教育支援システムの類似場面検索機能", 人工知能学会誌, Vol. 20, No. 2, 2005, pp.94-104.

田中 貴紘, 片上 大輔, 新田 克己 "調停トレーニングにおけるアドバイザーエージェント", 人工知能学会誌, Vol. 21, No. 4, 2006, pp. 319-329.

砂山 渡, 矢田 勝俊, "説得プロセス分析の枠組みと債権回収会話ログへの適用", 人工知能学会論文誌, Vol. 22, No. 2, 2007, pp239-247.

Y. Ohsawa and K. Yada (eds.), CRC Press, Data Mining for Design and Marketing, 2009, 319p.

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5件)

- ① W. Sunayama and K. Yada, "Analysis Framework for Knowledge Discovery related to Persuasion Process Conversation Logs", Data Mining for Design and Marketing, CRC Press, 2009, pp171-pp186, 査読なし
- ② K. Ohno and K. Yada, "Pricing System for Seeking Optimal Prices in the Diet Foods Market," Proc. of 2008 IEEE Conference on System, Man and Cybernetics, IEEE Press, 2008, pp. 3514-3518, 査読有.
- ③ 砂山 渡, 矢田 勝俊, "説得プロセス分析の枠組みと債権回収会話ログへの適用", 人工知能学会論文誌, Vol. 22, No. 2, 2007, pp239-pp247, 査読有.
- ④ K. Yada and K. Ohno, "Knowledge Discovery from Click Stream Data and

Effective site management", New Frontiers in Artificial Intelligence, Springer-Verlag, LNCS4384, 2007, pp360-pp373, 査読有.

- ⑤ K. Yada, CODIRO: A New System for obtaining Data Concerning Consumer Behavior Based on Data Factors of High Interest Determined by the Analyst, Soft Computing, Springer-Verlag, Vol. 11 No. 8, 2007, pp811-817, 査読有.

[学会発表] (計2件)

- ① 矢田勝俊, "Data Mining in Marketing", The Forth International Workshop on Data-mining and Statistical Science (DMSS2009), 2009年7月8日, 京大会館 (京都府)
- ② K. Yada, "Data Mining Application for Time-Series Data," Keynote Speech, Joint Conference on Information Sciences 2007 (July 20, 2007), 2007, Salt Lake City.

[図書] (計2件)

- ① Y. Ohsawa and K. Yada, CRC Press, Data Mining for Design and Marketing, 2009, 336p.
- ② 加藤直樹, 羽室行信, 矢田勝俊, 朝倉出版, 『データマイニングとその応用』, 2008, 194p.

[その他]

ホームページ等

6. 研究組織

(1) 研究代表者

矢田 勝俊 (YADA KATSUTOSHI)

関西大学・商学部・教授

研究者番号: 00298811