

人工知能学会 2015 年度研究会優秀賞 記念講演集

日時：2016 年 11 月 9 日(水)～ 12 日(土)

会場：慶応義塾大学 日吉キャンパス 来往舎

JSAI (一社) 人工知能学会

プログラム

11月9日(水) 13:10 - 14:30

[13:00-13:20] SIG-SLUD 言語・音声理解と対話処理研究会

1. 「説明行為の質の推定に向けた会話者のマルチモーダル情報モデリング」
○岡田将吾, 米航, 新田克己 (東京工業大学)
(SIG-SLUD-B501-07, 第 74 回, 2015 年 7 月 22 日) 04

[13:20-13:40] SIG-Challenge AI チャレンジ研究会

2. 「複数のマイクロホンアレイとロボット聴覚ソフトウェア HARK を用いた野鳥の観測精度の検討」
○松林 志保 (名古屋大学), 小島 諒介 (東京工業大学), 中臺 一博 (東京工業大学/
ホンダ・リサーチ・インスティテュート・ジャパン), 鈴木 霊璽 (名古屋大学)
(SIG-Challenge-043-11, 第 43 回, 2015 年 11 月 12 日) 10

[13:40-14:00] SIG-KST: 知識・技術・技能の伝承支援研究会

3. 「人の流れの計測とシミュレーションによる避難誘導方法の伝承
—新国立劇場における避難体験オペラコンサートを例に—」
○大西 正輝 (産業技術総合研究所), 山下倫央 (産業技術総合研究所),
星川 哲也 (新国立劇場), 佐藤和人 (新国立劇場)
(SIG-KST-026-06, 第 26 回, 2015 年 11 月 13 日) 16

[14:00-14:20] SIG-FPAI 人工知能基本問題研究会— 1

4. 「大規模グラフのコンパクトでスケーラブルな全距離スケッチ」
○秋葉 拓哉, 矢野洋祐 (国立情報学研究所)
(SIG-FPAI-B503-06, 第 100 回, 2016 年 3 月 27 日) 22

[14:20-14:40] SIG-FPAI 人工知能基本問題研究会— 2

5. 「高階結合正則化による時空間変化パターン検出」
○竹内 孝 (日本電信電話 (株)), 河原 吉伸 (大阪大学), 岩田 具治 (日本電信電話 (株))
(SIG-FPAI-B503-10, 第 100 回, 2016 年 3 月 27 日) 26

11月9日(水)発表分の受賞理由..... 32

11 月 10 日 (木) 13:10 - 14:50

[13:00-13:20] SIG-SWO セマンティックウェブとオントロジー研究会

6. 「記述論理に基づく農作業オントロジーの設計と応用」
○朱 成敏 (国立情報学研究所), 小出 誠二 (情報・システム研究機構),
武田 英明 (国立情報学研究所), 法隆 大輔 (農研機構), 竹崎 あかね (農研機構),
吉田 智一 (農研機構)
(SIG-SWO-038-06, 第 38 回, 2016 年 2 月 18 日) 34

[13:20-13:40] SIG-AM インタラクティブ情報アクセスと可視化マイニング研究会

7. 「情報アクセスにおける受動性と能動性: 音声対話によるニュース記事アクセス」
○林 良彦 (早稲田大学), 藤江 真也 (千葉工業大学/早稲田大学),
福岡 維新, 高津 弘明, 小林 哲則 (早稲田大学)
(SIG-AM-11-05, 第 11 回, 2015 年 11 月 13 日) 42

[13:40-14:00] SIG-KBS 知識ベースシステム研究会

8. 「アクティビティ評価情報と NMF の融合によるソーシャルメディアの信頼リンク予測」
○松谷 貫司 (龍谷大学), 熊野 雅仁 (龍谷大学), 木村 昌弘 (龍谷大学),
斉藤 和巳 (静岡県立大学), 大原 剛三 (青山学院大学), 元田 浩 (大阪大学)
(SIG-KBS-B502-02, 第 106 回, 2015 年 11 月 12 日) 50

[14:00-14:20] SIG-KSN 知識流通ネットワーク研究会

9. 「セルフサービスシステムの実現に向けた回答データ作成手法」
○竹内 広宜, 増田 聡本 晃太郎 (日本アイ・ビー・エム (株))
(SIG-KSN-018-06, 第 18 回, 2016 年 3 月 1 日) 58

11 月 10 日(木)発表分の受賞理由 64

11月12日(土) 13:10 - 14:30

[13:00-13:20] SIG-ALST: 先進的学習科学と工学研究会

10. 「Can LOD Based Question Generation Support Work in a Learning Environment for History Learning?」
○Jouault Corentin, Kazuhisa Seta, Yuki Hayashi (Osaka Prefecture University)
(SIG-ALST-B503-07, 第 76 回, 2016 年 3 月 6 日) 66

[13:20-13:40] SIG-LSE: ことば工学研究会

11. 「ニューラルネットワーク物語生成モデル」
○浅川 伸一 (東京女子大学)
(SIG-LSE, 第 49 回, 2016年 9 月 25 日) 72

[13:40-14:00] SIG-FIN 金融情報学研究会

12. 「株式掲示板におけるユーザ行動異常検知を用いた相場操縦発見手法に関する研究」
○宮崎 邦洋, 松尾 豊 (東京大学)
(SIG-FIN-015, 第 15 回 金融情報学研究会, 2015 年 9 月 25 日) 79

[14:00-14:20] SIG-DOCMAS データ指向構成マイニングとシミュレーション研究会

13. 「Twitter 位置情報・テキスト情報を用いた人の移動モデル構築と観光地推薦手法の提案」
○前田 高志ニコラス (東京大学), 吉田 光男 (豊橋技術科学大学),
鳥海 不二夫 (東京大学), 大橋 弘忠 (東京大学)
(SIG-DOCMAS, 人工知能学会合同研究会, 2015 年 11 月 14 日) 86

11月12日(土)発表分の受賞理由 94

今回発表されない受賞論文

- SIG-SAI 社会におけるAI研究会
「飲食店向け不動産営業を支援する申込み顧客推薦システムの評価」
河村 一輝, 諏訪 博彦 (奈良先端科学技術大学院大学), 小川 祐樹 (立命館大学),
荒川 豊, 安本 慶一 (奈良先端科学技術大学院大学)
(SIG-SAI-WSSIT, 第 25 回, 2016 年 3 月 1 日) 96

- SIG-BI: 経営課題にAI を! ビジネス・インフォマティクス研究会
「実データに基づく宅配便物流の地域集約方式による効率化の研究」
黒瀬 雄太 (東京工業大学), 寺野 隆雄 (東京工業大学)
(SIG-BI-004-01, 第 4 回, 2016 年 3 月 19 日) 104

今回発表されない受賞論文の受賞理由 110

説明行為の質の推定に向けた 会話者のマルチモーダル情報モデリング

A multimodal modeling for predicting the performance of storytelling

岡田 将吾^{1*} 米 航¹ 新田 克己¹
Shogo Okada¹ Mi Hang¹ Katsumi Nitta¹

¹ 東京工業大学大学院総合理工学研究科 知能システム科学専攻

¹ Dept. of Computational Intelligence and Systems Science, Tokyo Institute of Technology

Abstract: We present a multimodal analysis of storytelling performance in group conversation as evaluated by external observers. A new multimodal data corpus, including the performance score of participants, is collected through group storytelling task. We extract multimodal features regarding explanators and listener from a manual description of spoken dialog and from various nonverbal patterns. We also extract multimodal co-occurrence features, such as utterance of explainer overlapped with listener's back channel. In the experiment, we modeled the relationship between the performance indices and the multimodal features using machine learning techniques. Experimental results show that the highest accuracy is 82% for the total storytelling performance (sum of score of indices) obtained with a combination of verbal and nonverbal features in a binary classification task.

1 はじめに

対面会話において説明行為やストーリーテリングは情報共有, 他者への教示, などのために重要な役割を担う. 社会言語学の研究分野では, 説明における説得力を上げるためには, 説明中に用いられた言語情報だけでなく動作, 韻律, 表情といった非言語情報が重要であると報告されている [1]. 本研究は, 説明会話中の説明の質を, 会話中に交わされた言語・非言語情報から推定するモデルの構築に焦点を当てる. 外部観測可能な言語情報・非言語情報から会話中の説得力やストーリーテリングを推定する計算機モデルを構築出来れば, 説明会話の質を評価する技術や, 大規模な会話データから良質な説明シーンを検索するシステムに応用できる. またモデル構築・評価を通じて質の高い説明において観測されるマルチモーダル情報を分析することも可能となり, 説明行為の理解にも役立つ.

この目的に向けて, 複数のモーションセンサ・マイクを用いて説明会話を計測し, 説明会話において交わされた言語・非言語情報を含むマルチモーダルデータセットを新規に収集する. マルチモーダルデータセットは音声データ, 手・頭部の動作データといった複数の非言語データと, 発話内容を書き起こした言語データを含む. また, 収集した会話データを3名の観察者が閲覧し, 説明内容の基となった動画と見比べながら, 説明の円滑さ, 情報の正確さといった複数の観点から説明の質を評

価する.

次に, 発話区間, 韻律情報, 手のジェスチャ, 頭部のジェスチャ, 顔向け方向といった非言語特徴量を抽出し, 書き起こした言語データから発話内容中の単語特徴量を抽出する. さらに, 「共同注視をしながらの発話」といった会話における説明者と聞き手のインタラクションや, 「ジェスチャを伴いながらの発話」, 「聞き手を見ながらの発話」といったマルチモーダルパターンを共起パターンマイニングにより抽出する. 抽出した多様なマルチモーダル特徴量と, アノテーションされた説明行為の評価値の間の関係を分析, モデル化するために, 多変量解析, 機械学習を行い, 説明行為の質の推定精度について議論する. マルチモーダルモデリングの概要を図1に示す.

2 関連研究

対面会話において表出される参加者の音声区間, 韻律, 体・頭部の動作, 視線状態などの非言語情報に基づき, 参加者の役割 [2] [3], 性格特性 [4], リーダシップを有する人物 [5] といった参加者の高次特性を推定する研究が行われている. ただし, グループ会話における説明の質の推定に焦点を当てた研究は我々の知る限り存在しない. [6] はユーザが商品をのレビューを行っているビデオデータを収集し説得力の度合いをレビュアーの言語・非言語情報から推定する研究を行った. 本研究では説明者だけでなく, 聞き手とのインタラクションの過程で表出される言語・非言語情報を手掛かりに, 説明会話における説明の質を推定する.

本研究の貢献は以下の二つである. 説明の質を複数の

*連絡先: 東京工業大学 大学院総合理工学研究科
〒 226-8502 横浜市緑区長津田町 4259
E-mail: okada@dis.titech.ac.jp

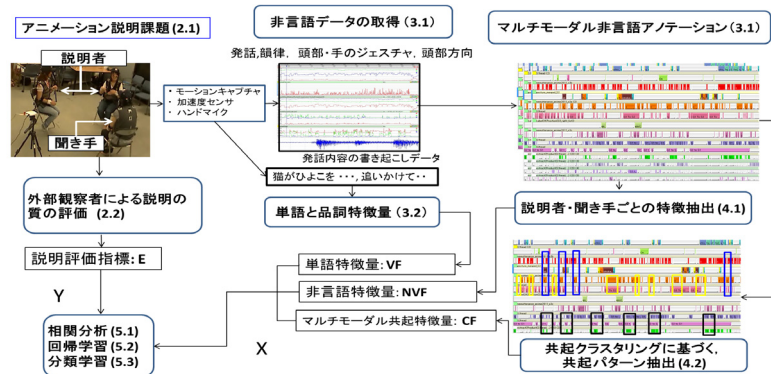


図 1: 説明評価値の推定のためのマルチモーダルモデリングの概要 (() 内の数字は論文中の節番号に対応している。)

外部観察者による説明への評価値として定義し、この評価値と、説明者・聞き手双方の言語・非言語情報とを関連付けることでグループ会話における説明の質を推定するモデルを構築・評価する。[2, 3, 4, 5, 6] では発話区間、ジェスチャ、言語といった個々の特徴量を 1 セッション中から抽出し、機械学習を行う際に統合していたが、複数のモダリティの時間共起特徴を考慮していなかった。本研究ではマルチモーダル共起パターンをデータマイニングのアプローチにより抽出し、モデルの推定に用いる。実験を通じて、説明の質を推定するために上記の共起特徴を含む多様な特徴量が説明の質を推定するために寄与することを示す。

3 データコーパス

3 人のグループ対話タスクとしてアニメーションの内容・各シーンの状況を説明する課題を設定した [7]。

3.1 会話タスクの概要

本研究では McNeill [8] により考案された、動画を事前に観察した参加者（以後、説明者と呼称）がその動画を見ていない参加者（以後、聞き手と呼称）に動画内容を説明するタスクを選定した。3 人は着座状態で対話を行う。人材派遣会社を通じて計 30 名の実験協力者を募集した。募集した 30 名はいずれも初対面の 20 代前半の女性であり、3 人同士を 1 グループとして説明タスクを 10 セッション行い、対話データを収集した。この内、十分に説明を行わなかった 1 グループと、センサデータの欠損が著しかった 1 グループのデータを除外し、計 8 セッションのデータを本研究に使用した。各セッションの平均対話時間は約 11 分（合計で約 700 分）であった。

アニメーションは 8 つのエピソードで構成されており、基本的には全てのセッションにおいて、この 8 つのエピソードに関する説明が行われる。また全てのセッションの説明者はアニメーションの状況設定、登場人物等の説明を事前に行ったため、事前説明を含め計 9 つのエピソードで構成される。本研究では説明会話をエピソードごとに予め分割し、8 セッションで計 67 エピソードの会話データセットを構築した。

3.2 外部観察者による説明の質の評価

本研究では、外部観察者により質の高い説明シーンを評価する。説明の理解度を基準とする場合、聞き手のインタビューに基づき説明の質を評価する必要がある。しかしながら、セッションごとに異なる聞き手が説明を受けており、インタビューの結果、聞き手の記憶量が一定でないことを確認したため、本研究では複数の第三者による評価値を採用した。説明の質の評価には、図 1 に示す 10 個の指標 (E_{1-10}) を用いた。 E_{1-10} は「雄弁性」、「熱意」、「円滑さ」、「機知に富んだ説明内容」、「情報の正確さ」、「端的な説明」、「説明内容の要約」、「活発な説明態度」、「協調的な説明」、「質疑応答の適切さ」に関する評価値をそれぞれ示す。最後に E_{Total} は説明評価値 E_{1-10} の総和を示しており、本研究ではこの値を総合的な説明評価値と定義する。スケールは 10 段階（最少 1、最大 10）とした。

上記の評価指標が妥当であるかどうかを判断するため 3 人の外部観察者による評価値の一致率をクローンバックの α 値により計算した結果、全て 0.7 を超えており、十分な一致率を得ることが出来た。各セッションの説明者は同一の動画内容に関する説明を行う。外部観察者はこの動画内容と、各セッションにおける説明内容を比較しながら説明の質を評価することができるため、十分な一致率を得ることが出来た考える。以降の実験では 3 人による評価値の平均値を、各指標の評価値として用いる。

4 マルチモーダルアノテーション

マルチモーダル情報の抽出のために、ハンドマイク、光学式モーションキャプチャ Mac3D、加速度・ジャイロセンサ (6 軸センサ) を用いた [7]。

4.1 非言語情報のアノテーション

音声区間検出

各話者の音声区間の検出を行う。ここでは零点交差法により音声区間の候補を抽出し、事前に音声区間を学習しておいた混合ガウシアンモデルを用いて音声区間を検

出した．上記の実行にはソフトウェア Julius¹ を用いた．音声区間として抽出された発話断片ラベル sp の集合を SP と定義する．

ジェスチャ区間検出

動作状態・無動作状態の 2 クラスに分類した．ジェスチャ区間の検出には正規分布を出力確率分布にもつエルゴディック型の HMM を用いた．手の動作区間として抽出されたハンドジェスチャラベル hg の集合を HG と定義する．

頭部ジェスチャアノテーション

後頭部に装着した 6 軸センサより取得される時系列データから頭部方向の動きを検出する．ここでは首を縦方向に動かしたか否かの 2 クラスに分類する．[9] で利用された，離散ウェーブレット変換により算出される特徴量を用い，識別器にはガウシアンカーネルに基づく SVM を用いてアノテーションを行う．頭部の動作区間として抽出されたジェスチャラベル he の集合を HE と定義する．

頭部方向アノテーション

本研究では顔向きから視線方向を近似する．三話者会話タスクにおける顔向きを分類するために，ある参加者が，参加者の左または右の参加者を見ているかどうかの 2 クラスに分類する．訓練データセットから前頭部・頭頂部マーカの位置座標の差分ベクトル集合を算出しておき，これらのベクトル集合と最近傍識別を行い，テストデータのラベルを分類した．参加者 i が参加者 l に視線を向けていると推定された区間の視線方向ラベル $g_{i,l}$ の集合を $G_{i,l}$ と定義する．

4.2 発話内容の言語アノテーション

説明の質は説明者のボキャブラリーに関連していると考えられる．各発言内容を日本語話し言葉コーパス (CSJ) の転記フォーマットに従って書き起こし，書き起こしたデータに対して形態素解析を行う事で発話内容を単語に分割し，品詞のアノテーションを行った．品詞の特徴量セットを W と定義し，各品詞を w と定義する．形態素解析には茶筌²を用いた．今回研究に用いた品詞セットは， w_1 ：名詞， w_2 ：接続詞， w_3 ：動詞， w_4 ：形容詞， w_5 ：副詞， w_6 ：前置詞， w_7 ：助動詞， w_8 ：感嘆詞の計 8 種類と w_9 ：フィラー， w_{10} ：笑い，のアノテーションを含め計 10 種類の特徴を用意した．

5 マルチモーダル特徴量の抽出

4 章でアノテーションしたマルチモーダルラベルセットから，説明の質を推定するために用いる特徴量の抽出を行う．まず説明者・聞き手ごとに発話量，発話内容に含まれる単語頻度といった個別の特徴量を抽出する．次に「ジェスチャを伴いながらの発言」といった各ラベル

表 1: 説明評価指標

説明評価指標	略記	質問事項
E_1 : (Eloquence)	Eloquent.	雄弁に説明していたか?
E_2 : (Enthusiasm)	Enthus.	説明者は熱意をもって説明していたか?
E_3 : (Fluency)	Fluent.	円滑な説明を行っていたか?
E_4 : (Wittiness)	Witt.	説明内容は機知に富んでいたか?
E_5 : (Preciseness)	Precise.	正確な情報を伝えていたか?
E_6 : (Compactness)	Compact.	簡潔に説明を行っていたか?
E_7 : (Summarization)	Summary	主要な点を要約して説明できていたか?
E_8 : (Liveliness)	Live	活気のある説明を行っていたか?
E_9 : (Cooperation)	Cooperate.	協調して説明を行っていたか?
E_{10} : (Q&A)	Q&A	質疑応答は適切であったか?
E_{Total}	Total	$E_1 - E_{10}$ の値の総和

を組み合わせた特徴量を抽出する．以下で述べる，マルチモーダル特徴量を表 2 にまとめる．全ての特徴量を，エピソードの時間長： EPL で割ることで正規化した．

5.1 各モダリティからの特徴量

ラベルセット SP, HG, HE, G に含まれる個々のラベルセグメントは始点と終点を有しており固有の時間長を有している． j 番目のエピソード SE_j にお

いて各参加者 i ($i = \{E1, E2, L\}$) から観測されるラベル系列から特徴量 $F_{i,j}$ を計算する．ここで $E1, E2$ は説明者， L は聞き手のインデックスを示す．各セッションの説明者は 2 名であるため，この 2 名から抽出した特徴量の和を説明者の特徴量として定義し，特徴量は各エピソード毎に計算される．

発話中の単語特徴量

j 番目のエピソード SE_j において各参加者 i の発話内容に含まれる品詞の回数 $W_{i,j}$ (4.2 節) を計算する．説明者の特徴量は $EW_j = W_{E1,j} + W_{E2,j}$ ，聞き手の特徴量は $LW_j = W_{L,j}$ として計算される． $W_{i,j}$ は 10 次元のベクトルであるため， EW_j, LW_j を合わせて 20 次元のベクトルが構成される．

発話ターンに関する特徴量

十分な情報量を相手に伝えるためには，一定の説明時間を確保する必要がある．発話の回数や発話量は説明の質に関連すると考えられる． j 番目のエピソード SE_j で参加者 i から N_i 個の発話断片が得られた場合の各発話断片を $sp_{i,k}$ ($1 \leq k \leq N_i$) として，各参加者の発話回数 SL_i ，発話長 ST_i を計算し，説明者の発話長 ESL ，発話回数 EST を以下のように計算する．

$$SL_i = \sum_k^{N_i} SL_{i,k}, \quad ST_i = N_i$$

$$ESL = SL_{E1} + SL_{E2}, \quad EST = ST_{E1} + ST_{E2} \quad (1)$$

聞き手に関するそれらの特徴量は $L_{SL} = SL_L, L_{ST} = ST_L$ とする．次に，会話中にターンテイキングが起きた回数を計算し，これを特徴量 TT として計算する．異なる話者の発話断片が順番に観測された場合に一回のターンテイキングとして認定する．この処理により発話ターンに関して 5 次元の特徴量を抽出した．

韻律に関する特徴量

発話中の声の抑揚は説得力を上げるために効果的であることが知られている [10]．このため本研究では発話

¹Julius: <http://julius.sourceforge.jp>

²Chasen: <http://chasen.naist.jp/hiki/ChaSen/>

断片中の音声データからエネルギーとピッチ（基本周波数）を計算する．ここでは発話者のみの特徴量を計算した．韻律情報の特徴抽出には Speech feature extraction code³を用いた．各セッション中に観測された2名の説明者の発話断片 $sp_{i,k}$ 群から計算したそれぞれのエネルギー E 、ピッチ P の値について最大、最少、平均、標準偏差をそれぞれ計算し、計8次元の特徴量を抽出した．

視線の特徴量

視線方向について、各参加者から見て左右の参加者のどちらを見ていたかのラベルが付与されている． k 番目の視線ラベルについて、参加者 l が参加者 i に視線を向けている時間を $GL_{i,l,k}$ と定義すると、説明者、聞き手が視線を向けられていた時間の総和 EGF 、 LGF は、 $GF_i = \sum_{l,k} GL_{i,l,k}$ 、 $EGF = GF_{E1} + GF_{E2}$ 、 $LGF = GF_L$ と計算される．次に、参加者 i, l の間で共同注視が観測された時間長を $MG_{i,l}$ と定義し、説明者同士の間での共同注視時間： $EMG = MG_{E1,E2}$ と説明者と聞き手の間での共同注視時間： $LMG = MG_{E1,L} + MG_{E2,L}$ を計算する．また、視線方向が変化した回数を TT と同様に計算し、説明者の視線変化回数： $EGT = GT_{E1} + GT_{E2}$ と聞き手の視線変化回数： $LGT = GT_L$ を計算した．視線について合計で6次元の特徴量を抽出した．

5.1.1 頭部と手のジェスチャの特徴量

頭部・手のジェスチャについても発話ターンと同様に、ジェスチャの持続長 HL_i, GL_i 、回数 HT_i, GT_i を各参加者 i ごとに計算し、説明者の持続長 EHL, EGL 、回数の総和 EHT, EGT を計算する．聞き手に関しても同様に計算し、頭部、手のジェスチャに関して $HEF = \{EHL, LHL, EHT, LHT\}$ の4次元、 $HGF = \{EGL, LGL, EGT, LGT\}$ の4次元の特徴量をそれぞれ抽出した．

5.2 非言語マルチモーダル共起特徴量の抽出

「視線を向けながらの発話」「ジェスチャを伴いながらの発話」といった、複数のモダリティの共起パターンは単一の特徴量よりも、詳細な会話者の会話態度や状態を示している．また、「説明者の発言時に聞き手がうなづいている」というような2者以上の間のインタラクションパターンも重要な会話のコンテキストを示していると考えられる．

本研究では、各参加者 i から取得したラベルセット $SP_i, HG_i, HE_i, G_{i,l}$ に含まれるセグメントの時間共起度に基づき、上記で述べるマルチモーダル共起特徴やインタラクションパターンを抽出する．本研究では効率良く時間共起度（同時に観測される割合）の高い共起パターンのみを抽出するために共起クラスタリングを利用する．本研究では [11] で提案されている共起クラスタリングを拡張して、この目的に利用する．

[11] のアルゴリズムでは、段階的に各パターン p （本研究では各ラベルに対応）をマージしていく戦略で、マ

表 2: マルチモーダル特徴量セット

	変数	特徴量の概要
VF	単語特徴量 (4.2 節)	
	EW_{1-10}	説明者の発話内容に含まれる品詞の頻度
	LW_{1-10}	聞き手の発話内容に含まれる品詞の頻度
	EPL	エピソードの時間長 (説明時間)
NVF	音声特徴量	
	ESL	説明者の発話長
	LSL	聞き手の発話長
	EST	説明者の発話回数
	LST	聞き手の発話回数
	TT	ターンテイキングの回数
	ESE	説明者の音声の中のエネルギー特徴
	ESP	説明者の音声の中のピッチ特徴
	視線特徴量	
	EGF	説明者が視線を向けられた時間長
	LGF	聞き手が視線を向けられた時間長
	EMG	説明者同士が共同注視した時間長
	LMG	説明者と聞き手が共同注視した時間長
EGT	説明者が視線方向を変化させた回数	
LGT	聞き手が視線方向を変化させた回数	
ハンドジェスチャの特徴量		
EGL	説明者がジェスチャを行った時間	
LGL	聞き手がジェスチャを行った時間	
EGT	説明者がジェスチャを行った回数	
LGT	聞き手がジェスチャを行った回数	
頭部ジェスチャの特徴量		
EHL	説明者がジェスチャを行った時間	
LHL	聞き手がジェスチャを行った時間	
EHT	説明者がジェスチャを行った回数	
LHT	聞き手がジェスチャを行った回数	
CF	マルチモーダル共起特徴量	
	CE_{1-195}	列挙された共起パターンの頻度

ジされた共起パターン $CP_x = \{p_1, p_2, \dots, p_{N_x}\}$ のみを候補として抽出していた．本研究では、各段階でマージされたパターンを全て列挙する．これは例えば「ジェスチャ(p_1)を伴いながらの発言 (p_2)」を示す共起パターン $\{p_1, p_2\}$ と「ジェスチャ(p_1)を伴いながらの発言 (p_2)中に聞き手がうなづいた (p_3)」を示す共起パターン $\{p_1, p_2, p_3\}$ とを別個のパターンとして抽出するためである．また、[11] のアルゴリズムではパターンの総数の多い順に共起パターン抽出を行っていたが、この場合、出現頻度の低いパターンを列挙できないため、本研究ではパターン総数の昇順、降順、両方の場合でアルゴリズムを実行し、列挙された共起パターン CE_x を特徴量として抽出する．この結果として計 195 の共起パターンが列挙された．共起パターン CE_x が各エピソード内で観測される回数を計算し、この回数を特徴量として定義する．紙面の都合でアルゴリズムの詳細については [11] に譲る．

6 マルチモーダル分析・モデリング

本章では、まず 6.1 節においてマルチモーダル特徴量と説明評価値との間の関係について相関分析し、説明評価値に対して有意に相関の高い特徴量を特定する．次に、6.2, 6.3 節では説明評価値の連続値を予測する回帰学習タスクと、評価値の高・低の2クラスを予測する分類学習タスクの2種類の機械学習タスクを行い、説明評価値の推定精度を検証する．

³Speech feature extraction code, <http://groupmedia.mit.edu/data.php>.

表 3: 説明評価値と有意な相関係数が得られたマルチモーダル特徴量 ($p < 0.05$: *, $p < 0.01$: **)

特徴量の 変数	最大の相関係数		有意な相関が 認められた評価指標
	評価指標	r	
EPL	E_8 (Live.)	+0.41**	E_{1-10}
単語特徴量			
EW(w_6)	E_8 (Live.)	-0.41**	E_{2-10}
EW(w_7)	E_1 (Eloquent)	-0.49**	E_{1-10}
EW(w_8)	E_1 (Eloquent)	+0.53**	E_{1-10}
EW(w_9)	E_{10} (Q&A)	-0.40**	E_{1-10}
EW(w_{10})	E_9 (Cooperate)	+0.27*	$E_{5,7,9}$
LW(w_3)	E_8 (Live.)	+0.30*	$E_{2,3,7,8}$
LW(w_4)	E_1 (Eloquent)	+0.31**	$E_{8,10}$
LW(w_7)	E_2 (Enthus)	-0.46**	E_{1-10}
LW(w_8)	E_2 (Enthus)	+0.32**	$E_{1-4,6,7,9,10}$
LW(w_9)	E_2 (Enthus)	+0.38**	E_{1-10}
音声特徴量			
ESL	E_1 (Eloquent)	+0.27*	$E_{1,3,4,8}$
LST	E_1 (Eloquent)	-0.27*	E_1
ESE _{min}	E_{10} (Q&A)	-0.49**	E_{1-10}
ESE _{std}	E_{10} (Q&A)	-0.28*	E_1
ESP _{min}	E_6 (Compact.)	-0.30*	E_{3-10}
ESP _{mean}	E_1 (Eloquent)	-0.30*	$E_{1-4,6,7}$
ESP _{std}	E_5 (Precise.)	+0.33**	$E_{2,4-7,9,10}$
視線特徴量			
EGF	E_5 (Precise.)	-0.27*	$E_{1,5,7}$
EMG	E_1 (Eloquent)	-0.35**	E_{1-10}
EGT	E_9 (Cooperate)	+0.45**	E_{1-10}

6.1 相関分析結果

本節では、表 2 に記載した全ての特徴量と 10 個の説明評価値の間でピアソン積率相関係数を計算し、 $p < 0.05$ の有意な相関値が得られた特徴量を報告する。表 3 は言語・非言語特徴量に関する相関分析の結果を示している。r は相関係数であり、 $p < 0.05$, $p < 0.01$ の有意な相関が認められた場合、相関値の横に *, ** をそれぞれ記載した。表 3 において、二、三列目には、一列目の特徴量との相関係数の絶対値が最大である評価値とその係数を、四列目には $p < 0.05$ で有意な相関を有する説明評価指標を記載する。

6.1.1 発話単語特徴に関する相関分析結果

表 3 より、説明者の発話内容に含まれる接続詞の頻度数 w_6 、参加者全員の助動詞の頻度数 w_7 は多くの説明評価指標と負の有意な相関を有している。この結果は説明の質が高いとされた説明シーン（エピソード）における説明者は接続詞や助動詞を多く使わないことを示す。一方で、簡単詞やあいづちが多く使われたエピソードの説明評価は高い傾向にあることがわかる。基本的にあいづちは重要な聞き手の行為であることが知られており、この結果は自然な結果であると考えられる。また説明者のフィラーの頻度 w_9 は説明評価値全てと負の相関が認められており、 E_1 と -0.40 の相関値を有することが示された。

6.1.2 非言語特徴に関する相関分析結果

表 3 より、エピソードの時間長 EPL は各説明評価指標と正の相関が認められることがわかる E_{1-10} (最大値は 0.41)。この結果は、質の高い説明をするためにはある

程度の時間をかける必要があることを示している。また、それと同時に、説明者の発話時間 (ESL) も説明指標と正の相関を有している。一方で聞き手の発話時間 (LSL) は、説明指標と負の相関を有している。聞き手がターンを長く保持するエピソードでは、説明の質が低い傾向にあることが示された。視線の特徴量において説明者に視線が集まる場合、また説明者同士が共同注視する時間 (EMG) が長い場合、説明の評価値が下がる (説明評価指標と負の相関にある) ことが示された。この原因をビデオを観察し分析したところ、説明内容に自信がない場面で、説明者同士が見合ってしまうケースが多いことがわかった。

6.2 回帰学習に基づく予測実験結果

本節では、抽出した特徴量セットが説明評価値の予測に有効かどうかを検証するため、回帰学習・モデル評価を行う。回帰モデルには L2 ノルムを正規化項に用いる、リッジ回帰モデルを採用した。交差検定法によりモデル評価を行う。ここでは 1 エピソードをテストとして用い、残りの 66 エピソードをモデルの学習に用いた。表 2 から 6 種類の特徴量セット (1)NVF, (2)VF, (3)CF, (4)VF+NVF, (5)VF+CF, (6)VF+NVF+CF (All) を用いて上記の評価実験を行い、有効な特徴量を検証した。マルチモーダル共起特徴とインタラクション特徴量 CE について、主成分分析を行う事により累積寄与率が 0.999 を超える最少数の主成分を特徴量として採用した。この結果、CE の次元数は 195 から 30 に削減された。交差検定法を行い、各試行でテストデータの予測誤差を算出し、全試行の予測誤差から決定係数 R^2 を計算した。リッジ回帰モデルの正則化項の重みパラメータを [1-150] と変化させ、訓練データ集合を使って最適なパラメータを求め、これをテスト時に使用した。図 2 は各特徴量セットを用いて学習した回帰モデルのテストデータに対する R^2 を各指標ごとに示している。

言語特徴 VF、非言語特徴 NVF、マルチモーダル共起特徴 CF で学習したモデルをそれぞれ比較すると、 $E_{1,5-8,10,Total}$ については言語特徴、 $E_{2-4,9}$ についてはマルチモーダル共起特徴を用いたモデルの精度が最も高かった。次に、全ての特徴量セットで学習したモデルの精度を比較すると、 E_7 については言語特徴が最大、 $E_{1,3-5,8,Total}$ については VF+CF: 言語とマルチモーダル共起特徴量で学習したモデルが最大、その他は全ての特徴量を用いて学習したモデルの精度が最大であった。

特に総合評価値 E_{Total} の精度は 0.24、精度の最大値は E_1 の評価値を予測した場合の 0.27 であった。テストデータに対する R^2 は最大で 0.3 未満とまだ十分ではない。一方で 67 エピソード全てのデータセットを使って、変数を p 値に従い追加しながら回帰モデルを学習し、 E_{Total} について自由度調整済みの R^2 を計算したところ 0.72 であった。この結果から、予測精度は充分ではないものの、提案した特徴量から全データセットを高

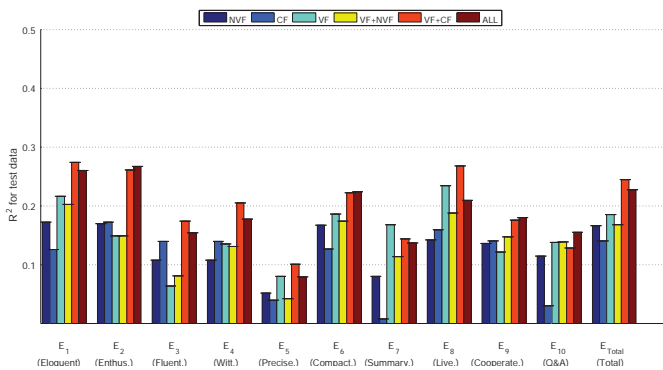


図 2: 説明評価値の回帰学習の結果

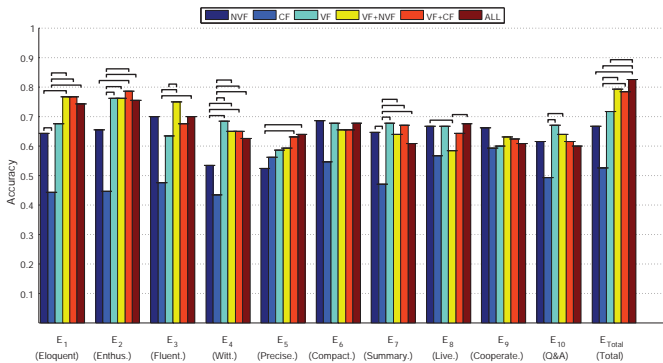


図 3: 説明評価値の高・低クラスを分類した結果 (□は t 検定の結果, 有意差 ($p < 0.05$) を認められたペアを示す.)

精度に説明するモデルを獲得できることが示された。

6.3 分類学習に基づく予測実験結果

中間値を閾値として用い、説明評価値を高レベルと低レベルの 2 クラスにカテゴリ化し、この高低レベルの分類タスクを行った。本タスクでは、訓練事例を増やすために、各エピソードをエピソード中の会話時間の中点を境に二分割した。この 2 分割されたサブエピソードには分割元のエピソードの評価値を割り当てた。言語・非言語・マルチモーダル共起特徴についても、サブエピソードごとに計算した。この操作により 134 個のサブエピソードデータセットを構築した。線形 SVM を用いて、10 分割交差検定を行った。訓練データセット内で交差検定を行い、SVM の C パラメータについて $[0, 0.01, 0.1, 1]$ から最適な値を選択し、テストに用いた。

6.2 節と同様に 6 種類の特徴量セットを用意し、上記の評価実験を行い、分類精度を比較した。図 3 は各説明評価値の高・低レベルを分類した結果を示す。縦軸はテストデータの平均正答率である。図 3 における □ は各モデル間の精度に対して t 検定を行った結果 $p < 0.05$ で有意差が認められたペアを示している。言語特徴 VF、非言語特徴 NVF、マルチモーダル共起特徴 CF で学習したモデルを比較すると、 $E_{2,4,7,10}$ について言語特徴で学習したモデルが有意に高い精度を得た。 $E_{1-3, Total}$ について、言語特徴 VF、非言語特徴 NVF、マルチモーダル共起特徴 CF いずれかを用いた場合よりも、特徴量を統合した場合の方が高い精度を得た。特に全ての特徴量で

学習したモデルは、総合的な説明評価指標 E_{Total} のレベルを 82.7% で分類出来ており、言語特徴を用いたモデルの 71.7%、非言語特徴量を用いたモデルの 66.6%、マルチモーダル共起特徴量を用いた 52.5% を上回った。この結果より、抽出した多様なマルチモーダル特徴量は外部観察者により評価された説明の質に関する指標を高精度に推定可能であることが示された。

7 結論

本研究は説明の質を外部観測可能な言語・非言語情報より推定するモデルの構築に焦点を当てた。この目的のために、グループ説明タスクを収録し、複数の外部観察者による説明の質の評価を行い、説明評価付きマルチモーダルコーパスを新規に構築した。発話内容から取得される単語の品詞特徴、発話区間、韻律、手・頭部のジェスチャ、視線方向といった多様な非言語特徴、マルチモーダル共起特徴・インタラクション特徴を加えたマルチモーダル特徴量を抽出することによって、説明評価値の推定を試みた。

実験の結果、各種推定精度については向上する余地はあるものの、多様なマルチモーダル情報から説明の質を推定することの可能性を示すことが出来た。今後の課題は、ビジネス、教育における説明場面など他の性質を持つ説明タスクを設定し、同様の実験を行い、説明の質を推定する上で、重要な不変特徴量とタスクに依存した特徴量を特定することである。

謝辞 本研究は科研費若手研究 (B) 25730132, 基盤研究 (B) 25280076, (C) 15K00300 の助成による。

参考文献

- [1] J. K. Burgoon, T. Birk, and M. Pfau, "Nonverbal behaviors, persuasion, and credibility," *Human Communication Research*, vol. 17, no. 1, pp. 140–169, 1990.
- [2] A. Vinciarelli, "Speakers role recognition in multiparty audio recordings using social network analysis and duration distribution modeling," *IEEE Trans. on Multimedia*, vol. 9, no. 6, pp. 1215–1226, 2007.
- [3] C.-Y. Weng, W.-T. Chu, and J.-L. Wu, "Rolenet: Movie analysis from the perspective of social networks," *IEEE Trans. on Multimedia*, vol. 11, no. 2, pp. 256–271, 2009.
- [4] F. Pianesi, N. Mana, A. Cappelletti, B. Lepri, and M. Zancanaro, "Multimodal recognition of personality traits in social interactions," in *Proc. of ICMI*, New York, NY, USA, 2008, pp. 53–60.
- [5] D. Sanchez-Cortes, O. Aran, M. S. Mast, and D. Gatica-Perez, "A nonverbal behavior approach to identify emergent leaders in small groups," *IEEE Trans. on Multimedia*, vol. 14, no. 3, pp. 816–832, 2012.
- [6] S. Park, H. S. Shim, M. Chatterjee, K. Sagae, and L.-P. Morency, "Computational analysis of persuasiveness in social multimedia: A novel dataset and multimodal prediction approach," in *Proc of ACM ICMI*, ser. ICMI '14, New York, NY, USA, 2014, pp. 50–57.
- [7] S. Okada, M. Bono, K. Takashi, Y. Sumi, and K. Nitta, "Context-based conversational hand gesture classification in narrative interaction," in *Proc of ACM ICMI*, 2013, pp. 303–310.
- [8] D. McNeill, *Hand and Mind: What Gestures Reveal about Thought*, ser. Psychology/cognitive science. University of Chicago Press, 1996.
- [9] K. Otsuka, H. Sawada, and J. Yamato, "Automatic inference of cross-modal nonverbal interactions in multiparty conversations: 'who responds to whom, when, and how?'" from gaze, head gestures, and utterances," in *Proc. of ACM ICMI*, 2007, pp. 255–262.
- [10] A. Tabensky, "Non-verbal resources and storytelling in second language classroom interaction," *Journal of Applied Linguistics*, vol. 5, no. 3, pp. 321–348, 2012.
- [11] A. Vahdatpour, N. Amini, and M. Sarrafzadeh, "Toward unsupervised activity discovery using multi-dimensional motif detection in time series," in *Proc. of IJCAI*, ser. IJCAI'09, 2009, pp. 1261–1266.

複数のマイクロフォンアレイとロボット聴覚ソフトウェア HARK を用いた 野鳥の位置観測精度の検討

Assessing the accuracy of bird localization derived from multiple microphone arrays and robot audition HARK

松林志保 鈴木麗凜

名古屋大学大学院情報科学研究科

小島諒介

東京工業大学大学院情報理工学科

中臺一博

(株) ホンダ・リサーチ・インスティテュート・ジャパン, 東京工業大学大学院
情報理工学科

要旨

本研究は、3つのマイクロフォンアレイとロボット聴覚ソフトウェア HARK を用いた野鳥の観測精度に関する予備的調査の結果を報告する。1つ目の調査では、愛知県豊田市の森林内でスピーカーからの鳥の歌の再生音を用いて HARK による定位精度の検討を行った。2つ目の調査では、人間による野鳥の観測結果と HARK により定位された野鳥の位置を比較し、その定位精度を調べた。

1 はじめに

近年、複数のマイクロフォンで構成されるデバイスであるマイクロフォンアレイを用いて音源の方向や位置を定位したり、定位した音源を分離する技術が急速に発展している。この技術の野鳥研究への応用は、従来の単一のマイクロフォンによる録音と比べてより豊富な生態情報の記録を可能にするため、生態理解へ大きな貢献を果たすことが期待される。

しかし、独自開発のシステムに基づく研究[1, 2]等ではなされているものの、この技術の野鳥研究への活用は未だ限定的な状況にあるといえる。その要因として、録音のためのデバイスの入手と分析のためのソフトウェアの利用が容易でない点や、野鳥の鳴き声の定位・分離性能の評価が十分でない点等が挙げられる。

我々はこれらの課題を克服すべく、ロボット聴覚オープンソースソフトウェアである HARK (Honda Research Institute Japan Audition for Robots with Kyoto University) [3]と、市販の会議用マイクロフォンアレイを用いたシステムを構築し、野鳥の歌に基づく方向の定位等を試みている[4]。HARK は、音響工学等に関する詳細な知識を必要とせず市販のマイクロフォンアレイ等を用いて PC 上に音源定位や分離等を含むシステムを柔軟に構築可能なソフトウェアであり、これまでロボットの聴覚のために開発されてきた。現在様々な文脈における応用が展開中であるが、野鳥の歌などの野外での音源定位の精度は十分

検証されていない。

本研究は、複数のマイクロフォンアレイと HARK を用いた野鳥の二次元位置推定精度を検討することを目的とする。その手段として2種類の実験を行った。まず、事前に録音された野鳥11種の歌を野外でスピーカーを用いて流し、マイクロフォンアレイからの距離や鳥の種類が HARK による定位精度におよぼす影響を測った。次に、HARK が実際のさえずりに基づいて定位した野鳥の位置を人間による野鳥観測結果と比較し、その定位精度を調べた。

2 手法

2.1. 複数のマイクロフォンアレイの設置

野鳥の録音実験は2015年6月、スピーカーテストは同年10月に名古屋大学フィールド科学教育研究センター稲武フィールド(愛知県豊田市稲武町)内の標高約1000m、樹齢60~70年のカラマツ植林と広葉樹の混合林内において、晴天ほぼ無風の気象条件下で行われた。録音には3つのマイクロフォンアレイ (Dev-audio 社 Microcone) を用いた。

各マイクロフォンアレイは、林内に定めた一辺が10mの正三角形の各頂点の位置に設置した三脚上に配置した。録音に用いた Microcone はそれぞれ7つ(水平方向に6つ、天井部分に一つ)のマイクロフォンから成る。

2.2 再生音源とスピーカーテスト

事前に録音された音源として、野鳥大鑑[5]付属のCD から調査地で営巣する野鳥11種の代表的なさえずりとその他の声(地鳴き等)を用いた(Table 1)。これらの音を、iPod と地上約1mに設置した外付けスピーカー (Sanwa supply bluetooth wireless speaker MM-SPBTBK) から一辺が10mの正三角形の中心、中心から西、北東、南東方向に25、50m離れた地点から流した(Figure 1)。25m地点でのマイクロフォンアレイの位置、スピーカーの距離および角度が正確に設置された場合に、鳥の歌の再生音が各

マイクロフォンアレイに届く理想的な角度を Table 2 にまとめる。

再生音は正三角形の中心に向けて流した。野鳥の歌の大きさは鳥の種類や競争相手の存在、外部音の有無などによる状況で異なるが、本実験では再生音の大きさは約35～40 dB の周辺音に対して約0～20 dB 大きい音（周辺音とほぼ同じかわずかに大きい音）で流した。音の大きさは無料の騒音測定器アプリ[6]を用いて測定した。同時に正三角形の頂点に設置した3つのマイクロフォンアレイを用いてスピーカーから流れる再生音を録音した。

Table 1 スピーカーテストに用いた野鳥リストと歌の種類。

鳥の種類	英名	鳥の名前コード	歌の種類
ウグイス	Japanese bush warbler	JBWA	さえずり、間奏
オオルリ	Blue-and-white flycatcher	BAWF	さえずり、間奏
ソウシチョウ	Red-billed leiothrix	RBLE	さえずり
キビタキ	Narcissus flycatcher	NAFL	さえずり、間奏
ヤマガラ	Varied tit	VATI	さえずり、威嚇音
ヒガラ	Coal tit	COTI	さえずり、地鳴き
センダイムシクイ	Eastern crowned willow warbler	ECWW	さえずり
ヒヨドリ	Brown-eared bulbul	BEBU	さえずり、時鳴き
カッコウ	Common cuckoo	COCU	さえずり
ツツドリ	Oriental cuckoo	ORCU	さえずり
ホトギス	Lesser cuckoo	LECU	さえずり

Figure 1 3つのマイクロフォンアレイと鳥の歌の再生音を流す地点の位置関係。一辺が10mの正三角形の北の頂点にマイクロフォンアレイ1、南の頂点にマイクロフォンアレイ2、東の頂点にマイクロフォンアレイ3を配置した。

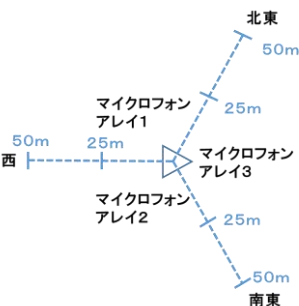


Table 2 三角形の中心、および各方向25m地点から鳥の歌の再生音を流した場合、音源が定位される理想的な角度。マイクロフォンアレイを設置した三角形の中心から北方向は0°、南方向は180°、東方向は-90°、西方向は90°とする。

	中心	西方向	北東方向	南東方向
マイクロフォンアレイ1	-150°	103°	-47°	-150°
マイクロフォンアレイ2	-30°	77°	-30°	-137°
マイクロフォンアレイ3	90°	90°	-17°	-163°

2.3 野鳥の位置観測と歌の分析

人間の観測者がマイクロフォンアレイを設置した正三角形の中心に立ち、録音開始と同時に周辺で観測された鳥の種類、中心からの大まかな位置等を約5分毎に記録した。野鳥の位置や種類は歌から推定し、正三角形の中心から25、50、100mの同心円を用いてフィールドノートに記録した[7]。野鳥

観測は目視と聞き取りに基づく (Figure 2)。



Figure 2 野鳥の観測と3つのマイクロフォンアレイを用いた録音風景。図中黄色い線は一辺が10mの正三角形を示す。

録音した野鳥の歌から、歌の再生と音声分析ソフトウェア Praat[8]を用いてのスペクトログラム（声紋）分析により手動で鳥種を分類し、個々の歌の始まりと終わりの時間を抽出した。これらの手動分析の結果を後述の HARK による音源定位結果と比較することで自動音源定位精度の検討を行った。

2.4 HARK による音源定位・分離・位置の推定

3つの各マイクロフォンアレイで収録した音声信号から方向・分離音を抽出するために、HARK を用いて音源定位・分離を行った。まずそれぞれのマイクロフォンアレイについて7チャンネルの音声信号を読み込み、短時間フーリエ変換によって得た各チャンネルのスペクトログラムから MUSIC 法[9]を用いて音源定位を行った。次にその定位結果をもとに Geometric High order Decorrelation based Source Separation(GHDSS)法[10]を用いて各音源方向に対応した音源を分離する音源分離を行った。

最後に、音源定位によって得た各マイクロフォンアレイを起点とした3つの方向（半直線）の交点を計算することで、音源の二次元空間内での位置を求めた。この時、音源定位の方向の誤差を許容するために、3つの半直線のすべての中心を音源とした3つの半直線のうちひとつでも交点を持たない半直線の組み合わせがある場合は誤検出として棄却した。

3 実験結果

3.1 スピーカーによる録音再生テスト

3.1.1. 定位音源の位置分布の確認

HARK により鳥の歌を自動定位した結果を参考にして、個々の分離音を人間の耳で確認することで再生音源との比較作業を行った。実際に再生音が定位された方角と、マイクロフォンアレイや音源の設置位置が理想的な場合の音源とマイクロフォンアレイの角度を Table 3 に示す。各方向毎に HARK が再生音を定位した角度と理想的な角度との差異に着目すると、西方向からの音源は北方向、北東方向からの音

源は南方向、南東方向からの音源は北方向と一定方向にずれが生じていた。このずれは HARK により定位された音源の 2 次元空間位置分布 (Figures 3~5) にも反映された。しかしながら Figures 3~5 に示されるように、システマティックな位置のずれはあるものの再生音はおおむね各スピーカーの位置付近で定位された。

再生音以外にマイクロフォンアレイ付近で定常的に定位された音源は、定位の際に 3 つのマイクロフォンアレイが異なる音源を同一の音源として定位したことにより生じたものと推測される。この現象は特に西、北東方向から再生音を流した場合に南東方向で頻繁に発生した (Figures 3 & 4)。これらの音源は南東方向に密生する笹群の葉音と推定される。逆に南東方向から再生音を流した場合には、北東方向にも再生音以外の音源が定位された (Figure 5)。北東方向の分離音を調べるとその多くは再生音の反響音であった。この反響音は付近のプレハブ小屋に起因すると考えられる。また、南東方向から再生音を流した地点は笹群内に位置する。笹群内では鳥の歌は定位されたものの分離しきれず、前後のさえずりや周囲の音と結合する現象も確認された。

Table 3 各マイクロフォンアレイで定位された実際の音源の方角と理想の方角の比較。音源とマイクロフォンアレイの位置関係は Figure 1 を参照のこと。マイクロフォンアレイを設置した三角形の中心から北方向は 0° 、南方向は 180° 、東方向は -90° 、西方向は 90° とする。

西方向(録音#137)	A: HARKが定位した方向	B: 理想の方向	AB間のずれ
マイクロフォンアレイ1	95°	103°	北方向に 8°
マイクロフォンアレイ2	55°	77°	北方向に 22°
マイクロフォンアレイ3	75°	90°	北方向に 15°
北東方向(録音#138)	A: HARKが定位した方向	B: 理想の方向	AB間のずれ
マイクロフォンアレイ1	-60°	-47°	南方向に 13°
マイクロフォンアレイ2	-45°	-30°	南方向に 15°
マイクロフォンアレイ3	-50°	-17°	南方向に 33°
南東方向(録音#139)	A: HARKが定位した方向	B: 理想の方向	AB間のずれ
マイクロフォンアレイ1	-145°	-150°	北方向に 5°
マイクロフォンアレイ2	-120°	-137°	北方向に 17°
マイクロフォンアレイ3	-150°	-163°	北方向に 13°

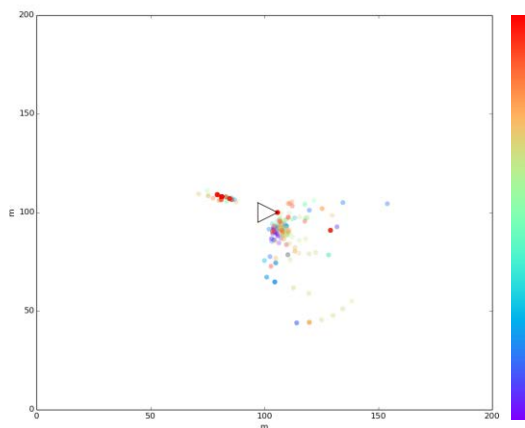


Figure 3 西方向 25 m 地点からの音源を定位した 2 次元位置分布。各色点は定位された音源を時系列で

示す。これらの音源は再生音以外の音源も含む。

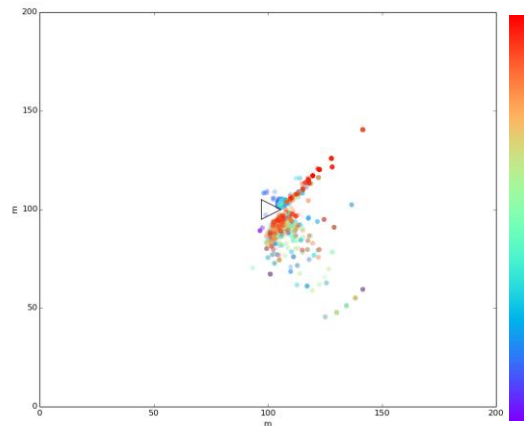


Figure 4 北東方向 25 m 地点からの音源を定位した 2 次元位置分布。各色点は定位された音源を時系列で示す。これらの音源は再生音以外の音源も含む。

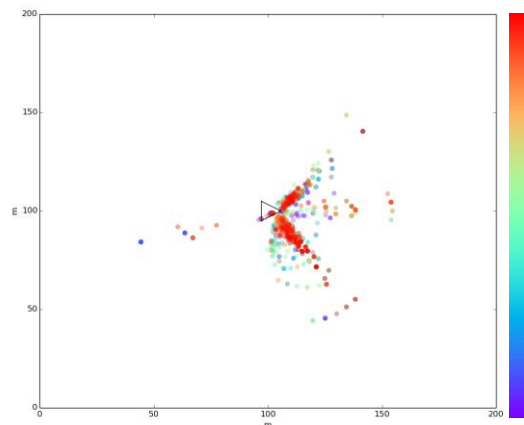


Figure 5 南東方向 25 m 地点からの音源を定位した 2 次元位置分布。各色点は定位された音源を時系列で示す。これらの音源は再生音以外の音源も含む。

3.1.2. 各方向における音源定位性能の検討

HARK による音源自動定位の精度は、音源からの距離と鳥の種類の影響を受けたようである。西、北東、南東方向における、25、50 m 離れた地点から鳥の歌の再生音流した場合の各方向の 3 つのマイクロフォンアレイの平均定位指数を Table 4 に示す。音源との距離が 25 m の場合は、3 方位とも全鳥種が定位され、全 11 鳥種の平均定位指数は各方向ともに 90 を超えた。一方、音源との距離が 50 m に伸びると、全鳥種の平均定位指数は西方向で 37.9、北東方向で 68.2、南東方向で 59.1 に低下した。

鳥種による違いに着目すると、50 m 地点における三方向の平均定位指数に見られるように、ウグイス、キビタキ、ヒガラなど比較的高音でさえずる種は 70 を超えたが、低周波の音域の歌を持つカッコウやツツドリは 20 以下となった (Table 4)。

Table 4 3つのマイクロフォンアレイによる平均定位指数。指数は、HARKにより分離された音を人間が耳で確認した際に、各マイクロフォンアレイが音源とほぼ同質の質を保ちつつ各鳥の歌を定位した場合にはそのマイクロフォンアレイに2、歌を部分的に定位した場合や分離精度が不十分な場合は1、全く定位しなかった場合は0のスコアを与えた後、3つのマイクロフォンアレイのスコアを合計した最大可能スコア（6）との比率を計算しその平均値を0から100までの値で正規化した。例えば、西方向50m地点での場合、ウグイスのスコアはマイクロフォンアレイ1では2、マイクロフォンアレイ2と3では各1ずつとなり、3つのマイクロフォンアレイの平均定位・分離指数は66.7となる。音源と3つのマイクロフォンアレイの位置関係はFigure 1を参照のこと。

	音源との距離25m				音源との距離50m			
	西方向	北東方向	南東方向	3方位平均	西方向	北東方向	南東方向	3方位平均
ウグイス	100	100	100	100	66.7	83.3	83.3	77.8
オオルリ	100	100	100	100	50.0	66.7	66.7	61.1
ソウシチョウ	100	100	100	100	33.3	83.3	83.3	66.7
キビタキ	100	100	100	100	66.7	83.3	83.3	77.8
ヤマガラ	100	100	100	100	33.3	83.3	50.0	55.6
ヒガラ	100	100	100	100	66.7	83.3	66.7	72.2
センダイムシクイ	100	100	100	100	33.3	66.7	50.0	50
ヒヨドリ	100	100	100	100	33.3	83.3	33.3	50
ツツドリ	33.3	66.7	100	66.7	0	0	33.3	11.1
カッコウ	100	100	100	100	0.0	33.3	16.7	16.7
ホトトギス	100	100	100	100	33.3	83.3	83.3	66.7
全11種平均	93.9	97.0	100	97.0	37.9	68.2	59.1	55.1

3.2. 野鳥の自動音源定位結果と人間による鳥観測結果の比較

Figure 6はHARKにより定位された音源の2次元位置分布を、Figure 7は人間の野鳥観測に基づく個々の野鳥の種類とその推定位置を示す。HARKによる音源の定位結果と人間の観測者による野鳥の推定位置を比較すると、その空間的分布パターンには類似性が見られた。

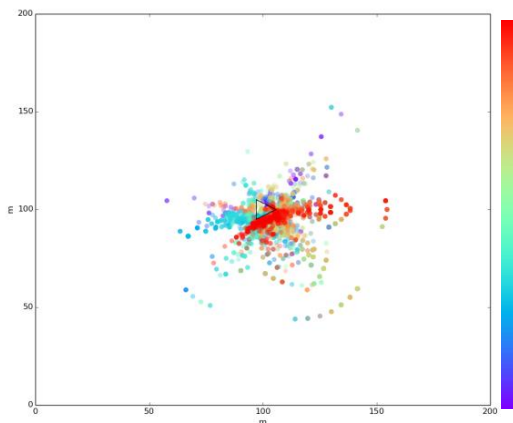


Figure 6 HARKにより定位された音源。各色点は定位された音源を時系列で示す。これらの音源は観測者の足音や周辺音等も含む。図中の正三角形は、マイクロフォンアレイを設置した正三角形に対応する。

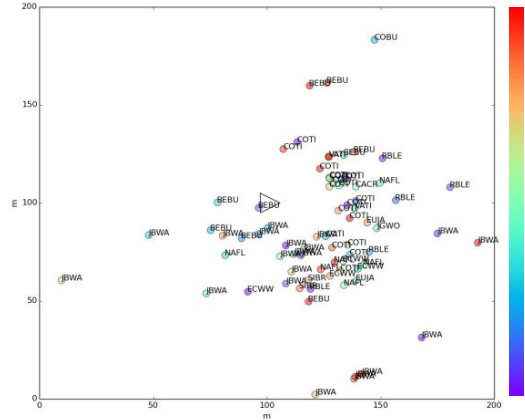


Figure 7 人間による野鳥観測結果。各色点は観測された鳥の種類と大まかな位置を時系列で示す。鳥の位置は16方位で表示した。鳥の名前コードはTable 1参照のこと。図中の正三角形は、マイクロフォンアレイを設置した正三角形に対応する。

HARKによる自動音源定位精度の検討のため、録音全体のスペクトログラムとその再生に基づく手動分析（歌の開始・終了時間の抽出と種のカテゴリ）と比較した。その一例をFigure 8に示す。この比較により、数個体の歌が一定方向で定位・分離されることを確認したが、その精度にはばらつきが見られた。

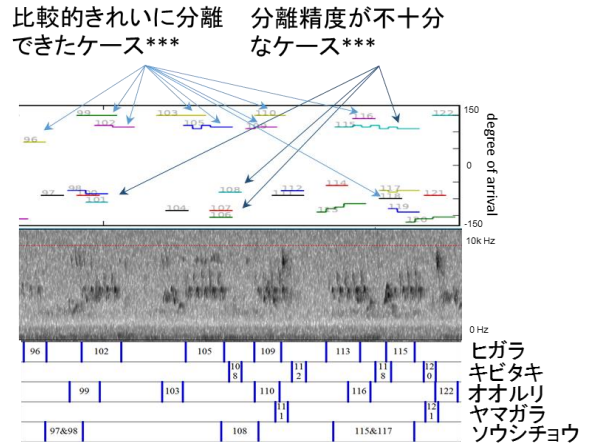


Figure 8 HARKによる鳥の自動音源定位・分離 vs. 人間による手動分析の一例。上段：HARKを使った自動音源定位。図中の各色線に対応する数字は分離されたファイル名を示す。中段：録音全体のスペクトログラム。数字は上記のファイル名に対応する。下段：スペクトログラムとその再生に基づく手動分析。人間の耳による分類が正しいという仮定に基づく。

精度にはばらつきが見られたものの、HARKによる音源定位結果は、人間による観測を補完する可能

性を示した。例えば、連続的なヤマガラの威嚇音にかき消され聞き落としていた他の種のさえずりがスペクトログラムにより明らかになった例や、さえずりの音質・バリエーションがよく似たキビタキとオオルリの判別に迷う際に、定位された鳥の位置を前後の時間帯の位置と比較することで区別が容易になった例などがある。

4 考察

スピーカーテストの結果、周辺音よりわずかに大きい音で流した野鳥 1 種の歌の定位距離の限界はおおよそ 50~75 m と推定された。この結果をもとに、HARK が定位した音源の二次元位置分布と人間による観測結果を比較すると、その分布パターンには類似性が見られた。この距離限界を超えると、人間の耳では容易に識別できる種、例えば比較的大きく特徴的な歌をもつウグイス (JBWA) やソウシチョウ (RBLE) も定位されなかった (Figures 6 & 7)。

定位限界距離の推定に加え、スピーカーテストは HARK による野鳥の定位精度 (accuracy) の検討には音源の位置だけではなく音源の分類作業が不可欠であることを明示した。その顕著な例として、3 方向の中で最も高い、位置の正確さ (precision) を示した西方向からの再生音実験 (Figure 3~5) が、最も低い平均定位指数 (Table 4) を示したケースが挙げられる。つまり西方向では定位された再生音自体が少なく、逆に北東および南東方向では比較的多くの再生音が定位されたが、定位された音源の中には再生音以外の多くの音源も含まれていた。

常時定位される野鳥の歌以外の音源はスピーカーテストだけではなく、HARK による定位結果と人間の観測者による野鳥観測の比較実験でも確認された。この現象は、各マイクロフォンアレイが異なる音源を定位しているにもかかわらず、同一の音源として処理することに起因する。これらの音源を除去するためには、ひとつひとつの分離音を人間が聞き分け各マイクロフォンアレイが同じ音源を指しているかを確認する作業が必要であるが、耳作業による音源の聞き分けは多大な労力と時間を要する。この事例は HARK による音源の自動分類性能の必要性を強く示唆する。

HARK による音源の定位性能は、音源からの距離だけではなく環境要因に大きく左右された。音が空気を振動して伝わる際には、空気そのものに加え、伝達途中にぶつかる障害物による減衰、吸収、拡散の影響を受けてひずみが生じる。音の伝達、ひいては HARK の定位性能に影響を及ぼした最も影響力の大きい障害物としては、録音現場付近のプレハブ小屋、地形、植生の 3 つが挙げられる。Figure 9 は、調査地周辺の航空写真と地形を示す。プレハブ小屋は北方向のマイクロフォンアレイ (Figure 1) にほぼ隣接し、北東方向からの音の伝達の障壁になっただけではなく、その他の方向からの音にも影響を及ぼした。小屋に起因する音の拡散や反響効果は、同じさ

えずり音が複数の方向で細切れに定位される事例や再生音以外の音源が一定方向で定位される事例 (Figure 5) から確認できた。

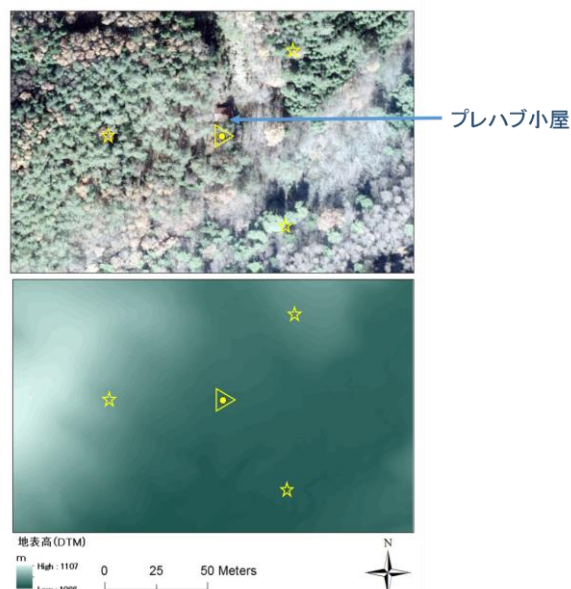


Figure 9 調査地周辺図。上: 航空写真 (撮影年不明、落葉期)。下: 航空機 LiDAR による数値地形モデル (2014年)。地形図に関する観測と作成は中日本航空株式会社による。図中の黄色い三角形は、3つのマイクロフォンアレイを設置した一辺が 10 m の正三角形に対応し、星印は各方向 50 m から再生音を流した地点を示す。

西、北東方向に広がる急勾配な地形 (Figure 9) も HARK による音源定位性能に影響を及ぼしたと考えられる。斜面に加え、音源とマイクロフォンアレイの間に位置する植生の影響も無視できない。特に比較的単純な森林構造を持つ針葉樹内に位置する西、北東方向から鳥の歌の再生音を流した場合、再生音以外の周辺音は主に南東方向で顕著に見られた。これは南東方向に位置する広葉樹林の林床に密生する笹群の葉音と推察される。笹の葉音は南東方向からの再生音を流した際には、HARK による分離性能を低下させ複数の音源を結合する現象を起こした。

音源からの距離に加えて、鳥の鳴き声の種類も定位精度に影響を与えた。一般的に低周波の歌は高周波の音に比べて減衰やそのほかの干渉を受けにくいため効率よく遠くに届く。また、森林への依存度が特に高い鳥は残響の影響や音のひずみを避けるため、比較的単純な構造の歌を歌うことが知られている [9]。本来であれば、スピーカーテストに使われた、ツツドリ、カッコウ、ホトトギス (いずれも Cuckoo 科) は順に約 500、800、1500 Hz [8] と比較的 low 周波かつ単純な構造の歌を歌うため、定位される可能性は高いことが予測された。しかしながら実際の平均定位指数は低い結果となった。これは、HARK による音限定の際にノイズカットのため 2000 Hz

以上に注目して定位を行ったため、特に低周波の音域で鳴くツツドリとカッコウの歌がカットされる結果となったためである。近距離で定位された音源は、これらの種の歌の一部のうち比較的高音部分がノイズカットをすり抜けた、もしくは偶然同方向の周辺音を拾った可能性がある。一方、高周波でさえずるヒガラや、比較的複雑な歌構造を持つキビタキは高い定位指数を示した。この一因としては、ヒガラやキビタキの歌の周波数が、風などの周辺の低周波の音とはっきり異なることがあげられる。今後のスピーカーテストで考慮すべき点として、音源の音質を鳥の周波数に絞ること、そして鳥種毎に歌の大きさを調整する必要がある。今回の実験では全鳥種を一定の地上高、大きさで流したが、実際には音の伝達効率には鳥の鳴く位置や環境の影響を受け、同時に鳥の歌の大きさは鳥の体の大きさなどに比例するからである。

5 おわりに

本稿は、複数のマイクロフォンアレイを用いた野外に置けるスピーカーテストおよび野鳥の音源定位精度の予備的調査の結果を報告した。スピーカーテストでは、晴天の無風状態で周辺音よりわずかに大きな音で鳥の歌の再生音を流した場合、その種類によりマイクロフォンアレイから約50～75mの距離まで定位できることが明らかになった。また、音源定位性能は周辺の人工物、地形、植生に加え、鳥の歌の周波数に影響を受けることが示された。これらの点を考慮した上で、HARKによる音源定位結果と人間による野鳥観測結果を比較すると相互間には類似した二次元位置分布が示された。さらに、分離された音源のスペクトログラムとその再生に基づく手動分析による種類の識別や歌の始まりと終わりの切り出しは、鳥がいつどこで鳴いたかという情報をより明確にした。このことは、HARKが人間による野鳥観測を補完する可能性を示唆している。

いずれの実験も初期的な段階にあるが、位置情報をもつ音声データを解析することは、野鳥の生態理解へ向け重要な意義を持つ。野鳥観測においては、瞬時に識別が難しい場合が頻繁に起こりうる。例えばオオルリやキビタキなど声の音質や歌のフレーズが似た個体が交互に鳴きその識別に迷う場合、さえずりの位置情報を前後の情報と比較することで2種の聞き分けが容易になった。このような例では特にデータの再現性が大きな意義を持つ。

今後の課題として、HARKによる自動分類機能の充実があげられる。野鳥のさえずりは、同種でも様々なレパートリーがあり、人間による手動分析は多大な時間と労力を要する。また、人的エラーの可能性も否めない。自動分類の機能が加われば分析の効率は格段に向上し、野鳥の位置的空間及び時間的空間利用の解明に向けた応用の可能性が高まると考えら

れる。

謝辞

高部直紀氏、近藤崇氏（名古屋大）のフィールドワークへのご協力に謝意を表す。また航空写真とLiDAR 地表モデルをご提供いただいた山本一清先生（名古屋大学）に感謝申し上げる。本研究の一部はJSPS 科研費 15K00335, 24220006 の助成を受けたものである。

参考文献

- [1] Collier, T.C., Kirschel, A.N.G., and Taylor, C.E. (2010). Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network. *Journal of Acoustical Society of America*, 128(1), 182-189.
- [2] Blumstein D.T., Mennill, D.J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J.L., Krakauer, A.H., Clark, C., Cortopassi, K.A., Hanser, S.F., McCowan, B., Ali, A.M., and Kirschel, A.N.G. (2011). Acoustic monitoring in terrestrial environmental using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48, 758-767.
- [3] Nakadai, K., Takahashi, T., Okuno, H.G., Nakajima, H., Hasegawa, Y., and Tsujino, H. (2010). Design and implementation of robot audition system "HARK"-open source software for listening to three simultaneous speakers. *Advanced Robotics*, 24, 739-761.
- [4] Suzuki, R., Hedley, R., and Cody, M.L. (2015). Exploring temporal sound-space partitioning in bird communities emerging from inter- and intra-specific variations in behavioral plasticity using a microphone array. Abstract book of Joint Meeting of the American Ornithologists' Union & Cooper Ornithological Society, 86.
- [5] 蒲谷鶴彦・松田道生著 日本野鳥大鑑鳴き声 420. (2011). 小学館.
- [6] Noise level meter. Retrieved October 1, 2015, from <https://itunes.apple.com/jp/app/noiselevelmeter/id694670057?ign-mpt=uo%3D5>
- [7] Ralph, C.J., Droege, S., and Sauer, J.R. Managing and monitoring point counts: standards and applications. (1995). USDA Forest Service general technical report. PSW-GTR 149.
- [8] Boersma, P and Weenink, D. (2015). Praat: doing phonetics by computer (Version 5.4.20) [Computer program]. Retrieved July 26, 2015, from <http://www.praat.org/>
- [9] Schmidt, R.O. (1986). Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions*, 34.3, 276-280.
- [10] Nakajima, H., Nakadai, K., Hasegawa, Y. and Tsujino, H. (2008). Adaptive Step-size Parameter Control for real World Blind Source Separation, In Proc. ICASSP 149-152.
- [11] Gill, F.B. (2007). *Ornithology*. NY: W.H. Freeman and Company.

人の流れの計測とシミュレーションによる避難誘導方法の伝承 — 新国立劇場における避難体験オペラコンサートを例に —

Transmission of Knowledge for Evacuation Drill using Pedestrian Tracking and Simulation — Example of Opera Concert with Evacuation Drill in New National Theatre, Tokyo —

大西 正輝^{1*} 山下 倫央¹ 星川 哲也² 佐藤 和人²

¹ 産業技術総合研究所 人間情報研究部門

² 新国立劇場運営財団

Abstract: Facility managers are required to perform an evacuation drill by Japanese law. However, it is difficult to perform the evacuation drill by many people. It is important to transmit the knowledge extracted from the evacuation drill definitely. In this paper we propose a transmission method of a knowledge for evacuation drill using pedestrian tracking and simulation. In the real experiment of the large-scale evacuation drill, the evacuated pedestrian flow was measured by this pedestrian tracking and evacuation time was analyzed by this simulation.

1 はじめに

消防法において施設の管理者は消防計画を作成し、定期的に消防訓練を行わなければならないことが規定されている。消防訓練とは消火訓練、避難訓練、通報訓練を指し、法令で定められていることから毎年、多くの施設では訓練が実施されている。例えば、総務省消防庁による避難訓練マニュアルでは火災発生時に火災発生場所を確認し、館内に通知した後、非常口や避難階段から避難誘導を行い、避難確認を行うまでの手順が示されている。しかし、日常業務の忙しさから多くの避難訓練はいわば「やりっぱなし」になっていることが多く、訓練から知見を抽出する作業や得られた知見を後々に効果的に引き継ぐ作業が十分に行われているとは言い難い。特に大規模な施設においては大人数の人員を動員しての避難訓練は容易ではない。そこで大規模な避難訓練を行う場合にはそこで得られた知見や知識を手軽に伝承するための技術の出現が期待されている。

不特定多数の人が出入りするような大規模空間における避難訓練から得られた知識や知見を正しく伝承するためには、避難訓練中の人の流れを正しく理解する技術 [1, 2, 3] が必要である。また、避難人数や避難経路などの条件を変えて何度も避難訓練を行うのは現実

的ではないことから、条件を変えた場合にどこで混雑が発生するかや避難時間がどのように変化するかをシミュレーションによって予測する技術 [4, 5, 6] が必要である。

筆者らはこれまでに北九州芸術劇場（2009年）と新国立劇場（2014年）において大規模な避難訓練を行い、その避難動線を計測すると共に、様々な条件における避難の様子をシミュレーションすることによってその後の避難誘導計画などに反映させるための研究を行ってきた。

本稿では、混雑した環境において人の流れを計測する技術とシミュレーションする技術について説明した後、2014年に新国立劇場で行った1300人規模の避難体験オペラコンサートについて説明し、これらの技術で解析することによって得られた知見を明らかにする。ここで得られた知見は新国立劇場だけではなく、他の大規模施設においても有用な知見であると考えられる。

2 人の流れの計測手法

ここでは、人の流れの計測手法 [2, 3] について簡単に説明する。人を計測するデバイスとしては RGB の色情報と深度 (Depth) 情報が同時に取得できる RGB-D カメラを用いる。RGB-D カメラの代表的なものとしてステレオカメラや Microsoft 社製 Kinect などがあげられる。Kinect は赤外パターンを投影するアクティブ

*連絡先：産業技術総合研究所 人間情報研究部門
〒305-8560 つくば市梅園 1-1-1 中央第1
E-mail: onishi@ni.aist.go.jp

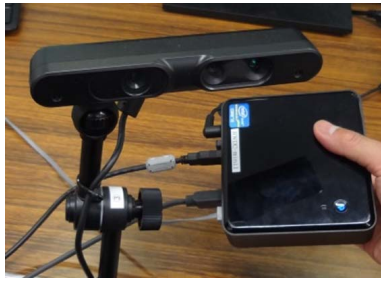


図 1: RGB-D カメラと小型 PC

ステレオ法によって深度画像を得ることができるカメラであり，室内空間においては高精度な三次元情報を得ることができる．本研究では Kinect と同じ原理である Xtion を用いて人の流れを計測する．図 1 に Xtion と得られたデータを処理する小型 PC を示す．Xtion は最大 8~10 [m] 程度の距離まで計測できることや電源が USB 供給であり配線が簡素であることなどから避難訓練の計測のような一時的な実証実験の場での利用に適している．

Xtion はおよそ 30 [フレーム/秒] で RGB-D 画像を取得可能であり，深度画像から 3 次元点群を復元することができる．復元した 3 次元点群から床平面を抽出した後，背景点を除去し，残った点群を二段階のファジークラスタリングをすることによって人の位置が計測可能である．本手法は深度画像を用いることで 2 次元画像からは抽出の難しいような混雑した環境においても正しく人の 3 次元位置を計測することができ，過去の人の位置を初期値として繰り返し演算に利用することで，連続した人の流れの抽出が可能である．これまでのステレオカメラを用いた実験では 0.56 [人/m²] 程度の混雑した環境においても 98.5 % 程度の精度で人の流れが抽出できており Xtion を用いた場合にはさらに高精度での抽出が可能であると期待できる．

3 人の流れのシミュレーション手法

次に，人の流れのシミュレーション手法 [5, 6] について簡単に説明する．人の流れのシミュレーションには主に避難行動を想定して作成したマルチエージェントシミュレータ CrowdWalk を用いる．CrowdWalk は一人ひとりのエージェント（人）に対して避難経路を設定することによって 1 ステップ毎の全体の避難状況をコンピュータで再現することができる．

CrowdWalk の特徴は避難経路を一次元で表現し，前方の人までの距離によって各エージェントの速度・加速度を決定することで処理を簡素化し，高速実行が可能な点であり，数千から数万，数十万の人の流れをシミュレーションすることができる．本シミュレータは

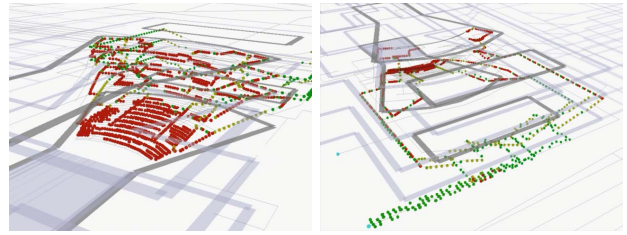


図 2: CrowdWalk によるシミュレーション例

北九州芸術劇場での 570 人の避難訓練を 5% 程度の誤差で再現することができている [7]．

図 2 に新国立劇場のオペラ劇場におけるシミュレーション結果を示す．小さな丸は一人ひとりのエージェント（人）を表しており，通常の場合は緑色，混雑して移動速度が減少している場合には黄色，ほとんど動けない場合には赤色で表現している．左図は避難が開始された段階であり，右図は左下に位置する避難場所に一部の人がたどり着いた段階である．

4 実証実験

4.1 避難体験オペラコンサート

2014 年 8 月 31 日に新国立劇場で避難体験オペラコンサートを開催した．本コンサートはオペラ鑑賞中に地震を原因とした火災が発生し，舞台責任者によって公演の中止が決定され，誘導員による誘導が行われるという想定でシナリオが進行する．観客は避難体験後に再開されるオペラを鑑賞することができる．なお，予め趣旨を伝えた上で体験希望者を募集した．地震が発生するタイミングについては教えていない．本コンサートの観客はおよそ 1300 人であり，劇場内の主要経路に 41 台の Xtion を設置し，避難の人の流れの様子を記録した．図 3 に地震が発生した直後のオペラ劇場の様子を示す．



図 3: 避難体験オペラコンサートでの地震直後の様子



(a) 1階 劇場内

(b) 1階 通路

(c) 2階 劇場出口

(d) 2階 出口

(e) 1階 出口 1

(f) 1階 出口 2

図 4: RGB-D カメラによる人の流れの計測結果

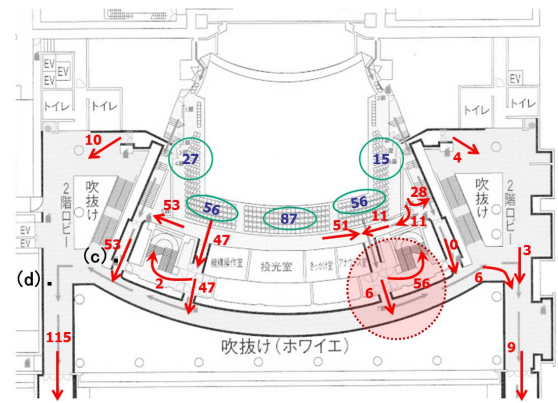
4.2 計測とシミュレーションの評価

図 4 に 41 台の RGB-D カメラによって抽出した人の流れの代表的な例を示す。また、フロアマップに通過人数を重畳した例を図 5 に示す。フロアマップの (a) ~ (f) は図 4 の (a) ~ (f) の映像が得られたカメラ設置位置に対応している。

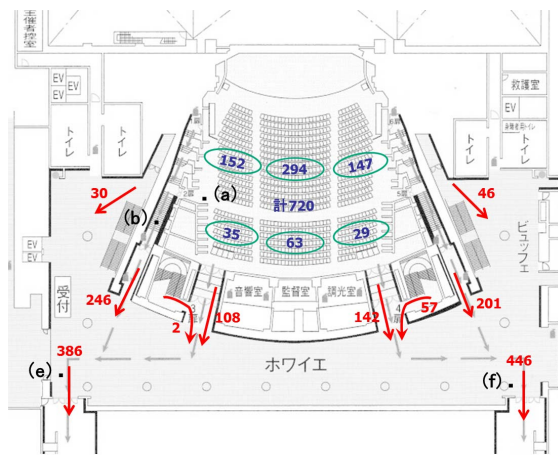
最も激しい混雑が発生した 1 階の出口 1 と出口 2 以外の計測場所の中から 100 人以上の人が通過した経路において本システムが計測した通過人数を目視によって評価した。表 1 に結果を示す。計測値と目視はそれぞれの方法で数えた通過人数を示しており、誤差は % で示している。平均誤差は 1.436% であった。なお、激しく混雑している出口 1 と出口 2 の評価については 4.3.1 で後述する。

表 1: 避難人数の計測結果と目視による評価

計測階	3F	3F	2F	2F	1F	1F	1F	1F
計測値	152	150	113	241	199	248	139	201
目視	158	150	115	246	201	246	142	201
誤差%	3.8	0	1.7	2.0	0.1	0.8	2.1	0



(a) 劇場 2 階の人の流れ



(b) 劇場 1 階の人の流れ

図 5: 新国立劇場にて観測された人の流れ

次に、シミュレーション結果について評価する。出口 1 と出口 2 にはそれぞれ 4 枚ずつ合計 8 枚の扉があり、例えば図 4 (e), (f) では共に 2 枚の扉が開放されている状態である。実際の避難訓練中にはこれらの扉の開放状態が時々刻々と変化するためシミュレーションでその様子を厳密に再現するのは難しい。

そこで、実際の避難訓練と人数や避難経路は同じ条件に設定し、最初から最後まで 1 枚の扉しか開かなかった場合と 4 枚の扉全てが開いていた場合をシミュレーションした。シミュレーションの結果を図 6 に示す。横軸が時間、縦軸が扉から出た人の合計人数（避難完了人数）を表している。最初の人の避難が完了してから最後の人の避難が完了するまでの時間は 4 枚全ての扉が開いていた場合には 194 秒、1 枚しか扉が開いていなかった場合には 509 秒かかっており、2 倍以上の時間が必要であることが分かった。実際に計測した避難訓練では 238 秒でありシミュレーション結果の扉 4 枚の場合と扉 1 枚の場合の間に収まっていることからシミュレーション結果は妥当であると考えられる。

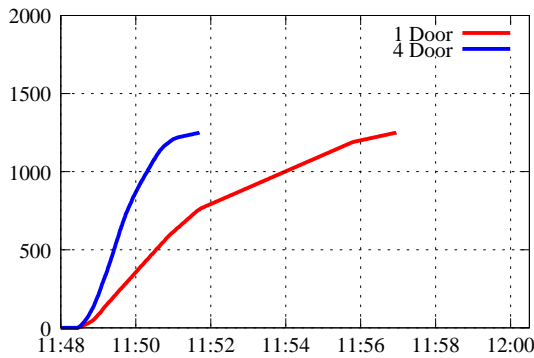


図 6: 扉の開放枚数が異なる場合のシミュレーション

4.3 実証実験の分析によって得られた知見

実証実験によって扉の開き方が避難に強く影響していることと一部の避難者が経路選択を誤った様子が確認できた。本節ではこの 2 つについて詳細に分析する。

4.3.1 扉の開閉状態の影響

火災が発生している場合、火が燃え広がるのを防ぐため扉を開放したままにしないことが求められている。一方で劇場では音漏れなどの防止のために扉は一般的には重い設計になっており閉まっている扉を開けるのは容易ではない。実際の避難訓練でも一度閉まった扉が容易に開かないという状態が散見された。誘導員が強制的に開いた場合にも、何かのはずみで再び閉まってしまうとすぐに開かれることはなかった。そこで本節では扉の開閉状態が避難時間にどのように影響するかについて検証した。

最も混雑した 1 階の出口 1 と出口 2 について検証した。図 7 にそれぞれの出口における時間と計測した単位時間あたりの通過人数のグラフを示す。出口 2 は出口 1 に比べて 4 枚の扉が効果的に使われていたため 60 人程度人数が多いにも関わらず、避難時間が 1 分程度短くなっていた。

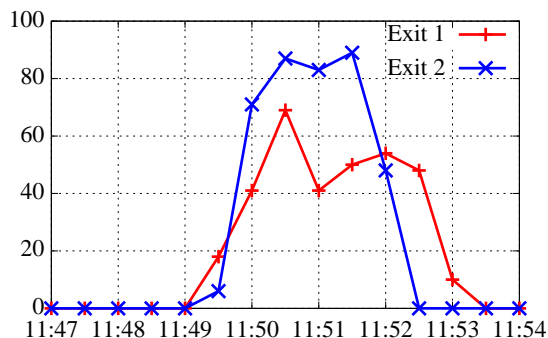
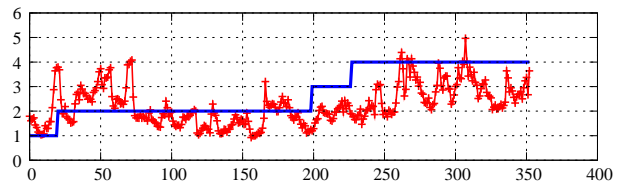


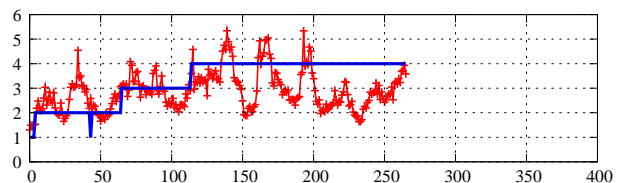
図 7: 出口 1 と出口 2 の通過人数 (30 秒毎)

表 2: 各出口からの避難人数と概算数

	出口 1	出口 2	合計
計測数	331	384	715
概算数	384	446	832
補正係数	1.166	1.161	1.164
取りこぼし率	14.249	13.901	14.063



(a) 出口 1 の扉の開放枚数と通過人数



(b) 出口 2 の扉の開放枚数と通過人数

図 8: 出口 1 と出口 2 の扉の開放枚数と通過人数

出口 1 と出口 2 は図 4 (e),(f) に示すように非常に混雑しているため目視によって人数を数えるのは困難であった。そこで周辺の通過人数から合計を求めたものを正解値として、抽出精度を評価する。表 2 に結果を示す。出口 1, 2 共に混雑が大きい時には 14%程度の人を取りこぼしているが、複数の地点で同程度の取りこぼし率であることから、現状の計測アルゴリズムは最も混雑している場合には 14% 程度の取りこぼしが発生するが 1.16 倍することで真値に近づくことが分かる。

次に訓練中に何枚の扉が避難に利用されているかを目視によって調べた。図 8 に結果を示す。青い線が避難時間に開いているドアの枚数を示し、赤い線が人の流れの計測技術で抽出した 1 秒あたりの通過人数を示している。出口 1 は 2 枚の扉が開かれている時間が長い一方で、出口 2 は早い段階で扉が 4 枚開いていることが分かる。

これらの関係から開放扉の枚数と単位時間あたりの通過人数の相関関係を求めた例を図 9 に示す。数字に多少の違いがあるものの、出口 1 (a) でも出口 2 (b) でも開放扉の枚数が多いほど、多くの人を通過させることができおり、これらの平均 (c) を取ると線形の関係に非常に近いことが分かる。扉が 4 枚使われた場合には 1 枚の時のおよそ 2 倍程度の人を通過させることができ、当然のことながら少しでも多くの扉を開放

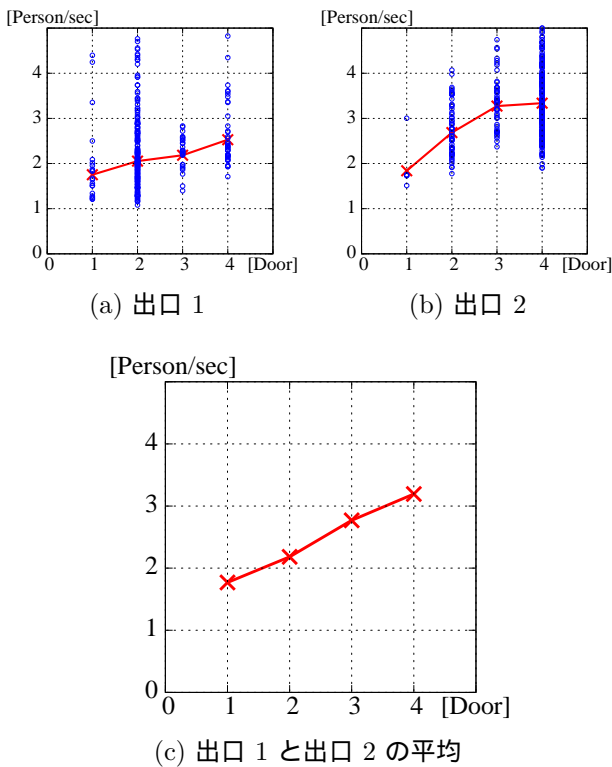


図 9: 扉の開放数と単位時間あたりの通過人数の相関

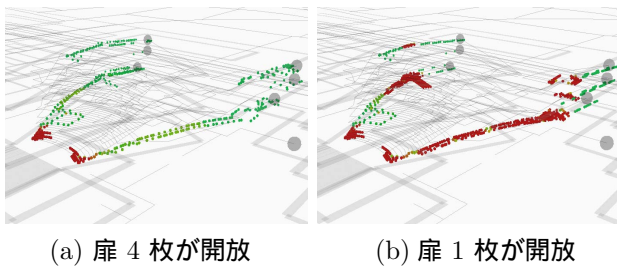


図 10: 同時刻のシミュレーション結果

することが避難時間の短縮には有効である。
 以上の結果はシミュレーション結果からも明らかである。扉の開放が 1 枚の場合と 4 枚の場合にどの程度の避難時間に差が出るかについては既に図 6 に示した。また、シミュレーション結果の俯瞰図を図 10 に示す。共に同じ時刻の状態を示しており左図は扉が 4 枚使われているのに対して右図は扉が 1 枚しか使われていない場合でありまだ多くの人々が劇場内に残り、続く廊下でも混雑が発生していることが分かる。

4.3.2 避難経路の選択誤りの影響

今回の避難訓練において先頭の集団が誤った経路を選択し、後方の人々が全てその人に付いていくという箇

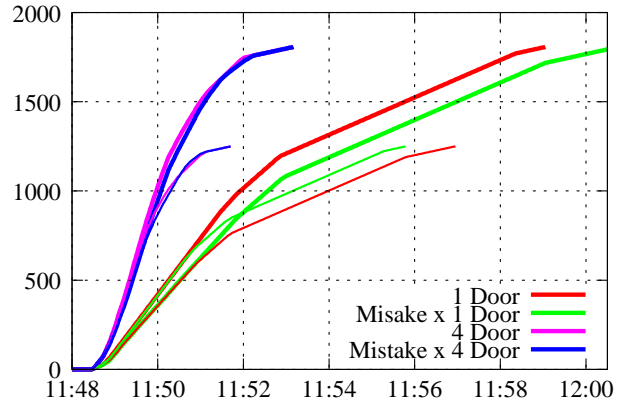


図 11: 様々な条件でのシミュレーション結果

所があった。図 5(a) の赤い網掛けが該当箇所である。その階に非常出口があるにも関わらず、先頭の人々が階段を降りてしまったため、後続の全ての人々が階段を降りてしまうという現象が見られた。

このような経路選択の誤りが全体の避難時間にどのように影響するかを検証するため観客が今回の避難訓練の 1300 人いた場合と満員となる 1800 人いた場合において以下の複数の避難条件を設定しシミュレーション実験を行った。

- (1) 誰も経路を間違えず出口の扉は 1 枚開放 (1 Door)
- (2) 避難訓練と同様に経路を誤り出口の扉は 1 枚開放 (Mistake × 1 Door)
- (3) 誰も経路を間違えず出口の扉は 4 枚開放 (4 Door)
- (4) 避難訓練と同様に経路を誤り出口の扉は 4 枚開放 (Mistake × 4 Door)

これらの組み合わせで合計 8 通りの条件でシミュレーションを行った。結果を図 11 に示す。横軸が時間を表し、縦軸が避難完了人数である。扉が 1 枚しか開放されていない場合は、そこで混雑が発生するため経路を誤った場合には、特に出口付近で混雑が激しくなり避難時間に影響を受ける。しかし、4 枚の全ての扉が開放されている場合には経路を誤ったとしても避難時間の遅れへの影響は極めて小さいことが分かる。

このように条件によっては 1 つのミスでは避難時間は大きく変わらないが、小さなミスが積み重なることによって避難時間が長くなってしまったり、少しでもミスをなくしていくことが避難時間の短縮には有効であることが伺われる。

訓練の後に行ったアンケートでは「何を手がかりに避難する方向を判断したか?(複数回答)」という質問に対して座席ごとに集計したところ表 3 のような結果が得られた。1 階中央、2 階上手のように係員の声が

表 3: 何を手がかりに避難する方向を判断したか？

	1 階中央	2 階下手	2 階上手	4 階上手
係員の声	20%	73%	25%	50%
係員の手振り	6%	12%	8%	15%
誘導表示	9%	13%	0%	3%
人の流れ	80%	39%	85%	65%
その他	4%	0%	13%	3%

聞こえにくかったところでは人の流れについて行く傾向がはっきりと出ている。特に 2 階上手は図 5 から分かるように誤った経路を選んだエリアであり、実際に先頭の人を選んだ誤った経路に付いて行っていることが観測されている。誘導には係員の声が大事であり、誘導が聞こえない場合には初期段階で正しい人の流れを作り出すことが大事であることが分かる。

5 むすび

本稿では大規模な人数を集めての避難訓練から知見を抽出するためのツールとして人の流れの計測手法と人の流れのシミュレーション方法について説明した。さらに、新国立劇場で行った 1300 人規模の避難体験オペラコンサートの避難動線の計測とシミュレーションを行い、扉の開放状態や経路の選択誤りがどのように避難に影響するかを明らかにした。このような大規模な人員を集めての実証実験は困難であり、実証実験から得られた知見や知識を伝承し、他の施設にも水平展開していく技術は重要な役割を果たす。

筆者らは施設だけではなく花火大会のような駅から街にかけての計測やシミュレーションも行っており [8]、オリンピックなどの大規模イベントを想定してより大規模な誘導支援へと展開していく必要があると考えている。

謝辞

本実証実験を行うために新国立劇場の関係スタッフの多大なる支援を得た。ここに感謝の意を表します。また、避難体験オペラコンサートに参加した 1300 人の参加者に心より感謝します。

参考文献

- [1] Alessandro Corbetta, Luca Bruno, Adrian Muntean, Federico Toschi, “High statistics measurements of pedestrian dynamics,” The Conference on Pedestrian and Evacuation Dynamics 2014 (PED2014).
- [2] 大西正輝, 依田育士, “ファジィクラスタリングを用いたステレオ画像からの動線抽出,” 電気学会論文誌, vol.128, no.9, pp.1438–1446, Sep. 2008 .
- [3] Masaki Onishi, “[Invited Paper] Analysis and Visualization of Large-Scale Pedestrian Flow in Normal and Disaster Situations,” ITE Transactions on Media Technology and Applications, vol. 3, no. 3 pp.170–183, July 2015.
- [4] Alessandro Pluchino, Cesare Garofalo, Giuseppe Inturric, Andrea Rapisarda and Matteo Ignaccolo, “Agent-Based Simulation of Pedestrian Behaviour in Closed Spaces: A Museum Case Study,” Journal of Artificial Societies and Social Simulation, vol.17, no.1, 16, 2014.
- [5] 山下倫央, 副田俊介, 大西正輝, 依田育士, 野田五十樹, “一次元歩行者モデルを用いた高速避難シミュレータの開発とその応用,” 情報処理学会論文誌, Vol.53 No.7, pp.1732–1744, July 2012 .
- [6] Tomohisa Yamashita, Shunsuke Soeda, Masaki Onishi, Itsuki Noda, “Exhaustive Testing of Evacuation Plan with High-Speed Evacuation Simulator,” International Scientific and Technical Conference Emergency Evacuation of People from Buildings, pp.357–364, March 2011.
- [7] 野中陽介, 大西正輝, 山下倫央, 岡田 崇, 島田敬士, 谷口倫一郎, “大規模な避難シミュレーションのための歩行速度モデルの精緻化,” 電気学会論文誌 C , vol.133, no.9, pp.1779–1786, Sep 2013.
- [8] 山下倫央, 大西正輝, “オリンピックにおける人の流れの解析,” 情報処理, vol.55, no.11, pp.1189–1195, Nov.2014 .

大規模グラフのコンパクトでスケーラブルな全距離スケッチ

Compressed All-Distances Sketches for Large Graphs

秋葉拓哉^{1*} 矢野洋祐^{1,2}
Takuya Akiba^{1,2} Yosuke Yano^{1,2}

¹ 国立情報学研究所

¹ National Institute of Informatics

² JST さきがけ

² JST, PRESTO

Abstract: The *all-distances sketch* is a useful neighborhood sampling scheme for large-scale graph analysis. In this study, we propose a compression technique to improve its space efficiency.

1 はじめに

All-distances sketches (ADS) [4, 5] は近年注目されるグラフに対するスケッチ手法である。全頂点に対する ADS はグラフサイズの線形に近い時間で構築を行うことができる。そして、ADS を用いると、近傍関数 [4, 9, 5], 距離 [6, 11], 近接中心性 [5], 近接類似性 [6], 実効直径 [2], 逆最近傍探索 [3], 時間を考慮した影響拡散 [10, 8, 7] といった、グラフ解析に用いる様々な指標が、効率的かつ高精度で推定可能である。

しかし、ADS は上述の通り理論的には優れた性質を持つものの、実際にはデータサイズが大きすぎ実用的ではないということが分かってきている。各頂点の ADS は長さが約 $k \ln n$ の配列である。ここで、 k は推定精度とサイズのトレードオフを設定するパラメータであり、 n は頂点数を表す。高精度な推定のためには、 k は数十から数百に設定されるため、ADS のサイズは対象のグラフ自体よりも大幅に大きなサイズになってしまう。

そこで、本研究では、新たなグラフのスケッチ手法である *sketch retrieval shortcuts (SRS)* を提案する。SRS は、ADS と類似した形式のデータ構造であるが、ADS よりサイズが小さい。そして、SRS から各頂点の ADS を、必要に応じて高速に復元できる。復元した ADS は、上述の様々な指標の推定に、通常の ADS と全く同様に利用できる。

構成 本論文の構成は以下の通りである。2 章で前提となる表記や概念を説明する。3 章で提案手法を説明する。4 章で実験結果を示す。5 章で結論を述べる。なお、本研究についての詳細は文献 [1] を参照されたい。

2 前準備

2.1 表記

$G = (V, E)$ を、頂点集合を V 、辺集合を E とする有向重みつきグラフとする。 $|V|$ と $|E|$ はそれぞれ n と m と表記する。辺の重みは 0 より大きいと仮定する。 $d(u, v)$ により頂点 u から v への距離を表す。 $P(u, v)$ は頂点 u から v への最短経路上の頂点を表す。即ち、 $P(u, v) = \{w \in V \mid d(u, w) + d(w, v) = d(u, v)\}$ である。

2.2 全距離スケッチ

All-distances sketches (ADS) は整数 k と頂点へのランク割当て r に基づき定義される。パラメータ k はデータサイズと推定精度のトレードオフを調整する。ランク割当て関数として $r: V \rightarrow [0, 1]$ を用いる。 $r(v) \sim U[0, 1]$ とする。即ち、 $r(v)$ は $[0, 1]$ の一様分布より選択される。

頂点 $u, v \in V$ について、 $N(u, v)$ を u に関して頂点 v よりも近い頂点の集合とする。頂点集合 $X \subseteq V$ について、 $k_r^{\text{th}}(X)$ を X 中で k 番目に小さいランク値とする。 $|X| < k$ である場合は $k_r^{\text{th}}(X) = 1$ とする。頂点 u, v について、 $\pi(u, v)$ を $\pi(u, v) = k_r^{\text{th}}(N(u, v))$ 、と定義する。ADS は以下のように定義される。

定義 1 (ADS [5]). 頂点 u の all-distances sketch (ADS) は $\mathcal{A}(u) = \{(v, \delta_{uv}) \mid v \in V, r(v) < \pi(u, v)\}$ である。ここで、 $\delta_{uv} = d(u, v)$ である。

$\mathcal{A}(u)$ の大きさの期待値は $O(k \log n)$ である [5]。

*連絡先：国立情報学研究所 東京都千代田区一ツ橋 2-1-2
E-mail: takiba@nii.ac.jp

3 提案手法

3.1 定義

提案手法 *Sketch retrieval shortcuts (SRS)* を定義する。ADS と同様に, SRS はグラフ $G = (V, E)$, パラメータ k , ランダムランク割当て $r: V \rightarrow [0, 1]$ に基づき定義される。

$\Delta = \{d(u, v) \mid u, v \in V\}$ とする。 $d_0 < d_1 < \dots < d_h$ とし $\Delta = \{d_0, d_1, \dots, d_h\}$ とおく。ここで, $d_0 = 0$ であり d_h はグラフの直径である。 \mathcal{B}_i ($i = 0, 1, \dots, h$), $\mathcal{C}_i, \mathcal{D}_i$ ($i = 1, 2, \dots, h$) を以下のように定義する。

- $\mathcal{B}_0(u) = \emptyset$ かつ $\mathcal{B}_i(u) = \mathcal{B}_{i-1}(u) \cup \mathcal{D}_i(u)$ ($i > 0$).
- $\mathcal{C}_i(u, v) = \{w \in P(u, v) \mid w \in \mathcal{A}(u), v \in \mathcal{B}_{i-1}(w)\}$.
- $\mathcal{D}_i(u) = \{(v, \delta_{uv}) \in \mathcal{A}(u) \mid \delta_{uv} = d_i, \mathcal{C}_i(u, v) = \emptyset\}$.

SRS は以下のように定義される。

定義 2 (SRS). 頂点 u の Sketch retrieval shortcuts (SRS) は $\mathcal{B}_h(u)$ である。

以下, 簡単のため, $\mathcal{B}_h(u)$ を $\mathcal{B}(u)$ と表記する。

定義より, $\mathcal{B}(u)$ は $\mathcal{A}(u)$ の部分集合である。従って, 大きさは以下のように評価できる。

補題 3. $\mathcal{B}(u)$ の大きさの期待値は $O(k \log n)$ である。

3.2 SRS からの ADS の取得

SRS の主な機能は, 任意の頂点の ADS を高速に再構築することである。SRS から ADS を取得することにより, 前述の通り, グラフ解析のための様々な指標が, 通常の ADS と同様に推定可能である。取得アルゴリズム Retrieve-ADS はアルゴリズム 1 である。頂点 u の ADS の取得は, SRS 上で $\mathcal{A}(u)$ に含まれる頂点のみに訪問するような最短経路探索に相当する。

アルゴリズム 1: 頂点 u の ADS の取得

```

Procedure Retrieve-ADS( $\mathcal{B}, u, k$ )
1   $A$   an empty all-distances sketch;
2   $Q$   a priority queue with only one element
   ( $0, u$ );
3  while  $Q$  is not empty do
4       $(\delta_{uv}, v) \leftarrow Q.$ Pop;
5      if  $u \notin A$  and  $r(v) < \pi(u, v)$  then
6          Add  $(v, \delta_{uv})$  to  $A$ ;
7          for all  $(\delta_{vw}, w) \in \mathcal{B}(v)$  do
8               $Q.$ Push( $\delta_{uv} + \delta_{vw}, w$ );
9  return  $A$ ;
    
```

アルゴリズム Retrieve-ADS の期待計算量は $O(k^2 \log^2 n \log(k \log n))$ 時間である。

3.3 ADS からの SRS の構築

まず, ADS から SRS を構築するアルゴリズムを説明する。詳細はアルゴリズム 2 として記述した。SRS は距離について再帰的に定義されているため, SRS の構築も距離が小さいエントリから順に行う。そのため, 全頂点の ADS の全エントリを, 距離が小さいものから順に, SRS に必要か否かを判断してゆく。エントリが SRS に必要か否かの判断に ADS 取得アルゴリズム Retrieve-ADS を利用する点がこのアルゴリズムの興味深い点である。

アルゴリズム 2: ADS からの SRS の構築

```

Procedure Construct-SRS( $G = (V, E), \mathcal{A}, k$ )
1   $B[u] \leftarrow \emptyset$  for all  $u \in V$ ;
2   $T \leftarrow \{(\delta_{uv}, v, u) \mid (v, \delta_{uv}) \in \mathcal{A}(u)\}$ ;
3  Sort  $T$ ;
4  for  $(\delta_{uv}, v, u) \in T$  do
5       $A \leftarrow$  Retrieve-ADS( $\mathcal{B}, u, k$ );
6      if  $(v, \delta_{uv}) \notin A$  then Add  $(v, \delta_{uv})$  to  $B[u]$ ;
7  return  $B$ ;
    
```

このアルゴリズムの期待計算量は $O(nk^3 \log^3 n \log(k \log n))$ 時間である。また, 毎回 Retrieve-ADS を呼び出す代わりに, SRS の各エントリの挿入時に, 取得可能な SRS エントリを陽に生成する工夫により, $O(nk \log n \log(nk \log n) + |\mathcal{B}| k \log n)$ 期待時間に改善する。ここで, $|\mathcal{B}|$ は SRS の合計サイズを表す。

3.4 直接的な SRS の構築

前述のアルゴリズム Construct-SRS は, 一度 ADS の陽な構築を経由してしまうため, 大きな作業領域をメモリ上に要求してしまう。そこで, 以下では, ADS を陽に構築することのないアルゴリズム Construct-SRS-Direct を考える。アルゴリズムの詳細はアルゴリズム 3 として記述した。ここでは, 簡単のため, グラフは重み無しを仮定する。このアルゴリズムは, 伝搬による ADS 構築アルゴリズムを仮想的に実行しながら, 前述の Retrieve-ADS によるエントリの要不要判定を行っている。

このアルゴリズム Construct-SRS-Direct の期待計算量は $O(D(n+m)k^2 \log^2 n \log(k \log n))$ 時間と $O(n+m+|\mathcal{B}|+k \log n)$ 空間である。ここで, D はグラフの直径を表す。ボトルネックとなる 7 行目の Retrieve-ADS の結果をキャッシュすることにより, 期待時間計算量は $O(Dnk^2 \log^2 n \log(k \log n) + mk \log n)$ に改善する。一方, 期待空間計算量は $O(n+m+|\mathcal{B}|+nk \log n)$ となるが, キャッシュするのは距離 $i-1$ のエントリのみでよく, 実験的には小さい。

アルゴリズム 3: SRS の直接的な構築

```
Procedure Construct-SRS-Direct( $G = (V, E), k$ )
1   $B[u] \leftarrow \emptyset$  for all  $u \in V$ ;
2  for  $i = 1, 2, \dots$  do
3     $f \leftarrow \text{FALSE}$ ;
4    for  $u \in V$  do
5       $T \leftarrow \emptyset$ ;
6      for  $v \in V$  such that  $(u, v) \in E$  do
7         $A \leftarrow \text{Retrieve-ADS}(B, v, k)$ ;
8        for  $(w, \delta_{vw}) \in A$  do
9          if  $\delta_{vw} = i - 1$  then Add  $w$  to  $T$ ;
10     Sort  $T$ ;
11      $A \leftarrow \text{Retrieve-ADS}(B, u, k)$ ;
12     for  $w \in T$  do
13       if  $r(w) \geq \pi(u, w)$  then continue;
14       if  $w \notin A$  then Add  $(w, i)$  to  $B[u]$ ;
15        $f \leftarrow \text{TRUE}$ ;
16   if  $f = \text{FALSE}$  then break;
17   return  $B$ ;
```

3.5 近傍の省略

SRS を用いてグラフを解析する際には、元のグラフ自体もメモリ上に存在しアクセス可能である場合が多いと考えられる。そのような場合には、そのグラフ自体の情報を併用することにより、SRS のサイズを更に減らすことが可能である。即ち、SRS に含まれるエントリの中から、元のグラフの辺に合致するものは削除できる。近傍を削除した場合、ADS 取得時の最短経路探索では、SRS のエントリだけでなく元のグラフの辺も遷移に用いる。

4 評価実験

4.1 実験方法

本実験には CPU が Intel Xeon 2.67 GHz 2 ソケット、メモリが 96GB の Linux サーバを利用した。全てのアルゴリズムは C++ で実装され、gcc 4.8.4 を用いて最適化オプション `-O3` の設定でコンパイルされた。データセットとして実際のソーシャルグラフとウェブグラフを用いた (表 1)。SRS 構築アルゴリズムは並列化され 24 スレッドで実行した。 $k = 16$ とした。

手法として以下の 5 つを比較する。(1) ADS は通常の ADS である。(2) ADS-c は LZ 系圧縮アルゴリズム (google-snappy) を適用した ADS である。(3) SRS は ADS 経由で SRS を構築する手法である (アルゴリズム 2)。(4) SRS-d は SRS の直接的な構築アルゴリズムである (アルゴリズム 3)。(5) SRS-i は 3.5 章で述べた近傍の省略を用いた SRS である。

表 1: 実験で用いたデータセット

名称	種別	$ V $	$ E $
email-Enron	Social (u)	36,692	367,662
com-dblp	Social (u)	317,080	2,099,732
web-Google	Web (d)	875,713	5,105,039
in-2004	Web (d)	1,382,870	16,917,053
flickr-links	Social (u)	1,715,256	31,101,563

4.2 実験結果

表 2 が実験結果を表す。

スケッチサイズ ADS-c が改善を達成していないことから、一般的な圧縮アルゴリズムはあまり効果的でないことが確認できる。一方、SRS は ADS に対し大幅な改善を達成している。

構築の時間と空間 SRS は ADS より構築に時間が掛かる。SRS-d は更に時間がかかる。一方、一般に SRS-d が最も小さな空間で構築を達成する。SRS の所要空間は ADS に比例するが効率的な操作のためのデータ構造により数倍大きくなってしまっている。

取得時間 SRS による ADS の平均取得時間は概ね 1 ミリ秒以下であり非常に高速である。SRS-i は SRS と比較すると低速になるものの数ミリ秒程度であり許容範囲であると考えられる。

5 おわりに

本研究では、all-distances sketches (ADS) の高速な取得のためのデータ構造 sketch retrieval shortcut (SRS) を提案した。各頂点の ADS は SRS より高速に復元でき、通常の ADS と同様に、様々なグラフ解析の指標の推定に利用可能である。

謝辞

本研究は日本学術振興会科学研究費補助金 (15H06828) 及び JST さきがけの支援を受けたものである。ここに記して謝意を表す。

参考文献

- [1] T. Akiba and Y. Yano. Compact and scalable graph neighborhood sketching. Manuscript, 2016.
- [2] P. Boldi, M. Rosa, and S. Vigna. HyperANF: Approximating the neighbourhood function of very large graphs on a budget. In *WWW*, pp. 625–634, 2011.

表 2: 左側は, スケッチのサイズと ADS の平均取得時間を表す. 右側は, 構築時の所要時間と所要空間を表す.

データセット	サイズ (MB)				構築空間 (MB)			
	ADS	ADS-c	SRS	SRS-i	ADS	ADS-c	SRS	SRS-d
email-Enron	19.46	19.63	1.53	0.56	59.11	59.40	182.85	40.57
com-dblp	222.15	223.73	22.30	14.66	529.67	532.26	1929.51	303.22
web-Google	451.38	455.01	78.42	58.45	1055.01	1052.66	3956.93	734.14
in-2004	597.66	603.18	138.63	92.68	1468.30	1489.94	5049.08	1272.42
flickr-links	1277.11	1285.42	59.56	16.84	2866.05	2893.68	11344.86	2045.52
	取得時間 (μ s)				構築時間 (s)			
	ADS	ADS-c	SRS	SRS-i	ADS	ADS-c	SRS	SRS-d
email-Enron	—	0.84	182.71	538.69	2.26	2.40	5.32	11.46
com-dblp	—	1.33	413.76	494.32	45.83	46.12	84.60	324.12
web-Google	—	1.16	271.13	260.74	82.57	81.32	173.13	952.80
in-2004	—	0.91	160.90	189.82	64.67	66.12	171.73	3473.29
flickr-links	—	1.85	517.60	8163.60	341.40	319.07	599.38	1954.50

- [3] E. Buchnik and E. Cohen. Reverse ranking by graph structure: Model and scalable algorithms. *CoRR*, abs/1506.02386, 2015.
- [4] E. Cohen. Size-estimation framework with applications to transitive closure and reachability. *J. Comput. Syst. Sci.*, 55(3):441–453, 1997.
- [5] E. Cohen. All-distances sketches, revisited: HIP estimators for massive graphs analysis. *IEEE TKDE*, 27(9):2320–2334, 2015.
- [6] E. Cohen, D. Delling, F. Fuchs, A. V. Goldberg, M. Goldszmidt, and R. F. Werneck. Scalable similarity estimation in social networks: closeness, node labels, and random edge lengths. In *COSN*, pp. 131–142, 2013.
- [7] E. Cohen, D. Delling, T. Pajor, and R. F. Werneck. Sketch-based influence maximization and computation: Scaling up with guarantees. In *CIKM*, pp. 629–638, 2014.
- [8] E. Cohen, D. Delling, T. Pajor, and R. F. Werneck. Timed influence: Computation and maximization. *CoRR*, abs/1410.6976, 2014.
- [9] E. Cohen and H. Kaplan. Summarizing data using bottom-k sketches. In *PODC*, pp. 225–234, 2007.
- [10] N. Du, L. Song, M. Gomez-Rodriguez, and H. Zha. Scalable influence estimation in continuous-time diffusion networks. In *NIPS*, pp. 3147–3155, 2013.
- [11] M. Thorup and U. Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005.

高階結合正則化による時空間変化パターン検出

Structured regularizer for spatio-temporal matrix completion

竹内 孝^{1*} 河原 吉伸² 岩田 具治¹

Koh Takeuchi¹ Yoshinobu Kawahara² Tomoharu Iwata¹

¹ NTT コミュニケーション科学基礎研究所 ² 大阪大学産業科学研究所

¹ NTT Communication Science Laboratories ² ISIR, Osaka University

Abstract: We propose a structured regularizer for a problem of completing missing values in a matrix. We apply our regularizer for monthly recorded weather data and show extracted patterns.

1 はじめに

近年、様々な現象の観測を目的とした多種類センサーの設置が行われており、高精細な時空間解像度を持つ多変量時系列データが取得されるようになった。このようなデータから変化点を検出する技術は、都市交通流解析、不正アクセス解析、バイオインフォマティクス、医療情報学、気象学、公衆衛生学、公衆安全学など多数のアプリケーションから注目を集めている [10, 18]。公衆衛生学では、感染症のアウトブレイクが、いつどこで発生したかを変化点検出によって特定し、将来起こりうるアウトブレイクの予防に役立てられている [18]。また、交通流解析では都市の交通状況がいつどこで変化したかを検出することで、都市計画を立てる際に交通渋滞や大気汚染削減のための参考として利用されている [18]。

時空間データの特徴量は、センサーの位置と観測時刻からなり、特徴量そのものが補助的な構造を持つ。例えばコンピュータネットワークから取得されるデータでは、コンピュータがセンサーに対応するため、コンピュータ間の距離やネットワークのトポロジー、同一タスクの並列処理を行うコンピュータのグループなどの補助的な構造が存在する。また時刻に関しても、複数のセンサーが一定期間や、特定の曜日に同一の値を取るなどのダイナミクスや周期的な現象などが存在する。特徴量に関する補助構造は、時空間データの変化を捉えるために有用な事前情報になると考えられる。

本研究では、変化点検出問題を一般化し、欠損値を含む多次元時系列データから時間と空間の両方向に起きる変化パターンを前述の補助構造を活用して検出する手法を提案する。ここでは時空間変化パターンとは、平均値変化 (mean-shift)、すなわち観測値の平均値がある値から異なる値へ移動する箇所を特定する問題と

する。例えば時間方向の変化だけに注目すると、2つの変化点の間に存在する時刻は同一クラスターに所属していると考えることで、平均値変化点検出問題は時系列データの分割問題として扱うことが出来る。近年、構造正則化を用いた時間変化点と分割毎の平均値を同時に推定する手法が提案されている [2, 11]。しかし、これらのモデルでは時間の連続性しか構造正則化に用いていない。さらに観測機器の故障やネットワーク転送の失敗などにより生じる欠損値を考慮していないため、本研究の問題には適応できない。

そこで本研究では、既存の構造正則化の一般化として、時間と空間に関する複数種類の事前情報を単一の枠組みで柔軟に利用できる構造正則化を提案する。さらに正則化項が、劣モジュラ関数のロバッシュ拡張から得られる凸関数である性質を利用し、正則化項を最小化する近接法を提案する。データの特徴量を d とすると、我々の提案する近接法は、最大流アルゴリズムを利用することで正則化項の最小化問題の厳密解を近似的には $O(d \log d)$ の計算コストで得ることが出来る。本研究では、提案した構造正則化を行列補完問題と組み合わせることで、欠損値を含む時系列データからの変化パターン検出を行う。

2 結合正則化による変化点検出

ある位置 i_1 に設置したセンサーから時刻 i_2 に観測された値 x_{i_1, i_2} からなる行列を $X \in \mathbb{R}^{I_1 \times I_2}$ とする。 X にはノイズが加算されていると仮定し、観測値の潜在平均値からなるパラメータ行列を $Z \in \mathbb{R}^{I_1 \times I_2}$ とする。前後する時刻 i_2 と $i_2 + 1$ において、 $z_{i_1, i_2} = z_{i_1, i_2 + 1}$ ならば、 (i_1, i_2) と $(i_1, i_2 + 1)$ は同一のクラスターに属するとみなす。

結合正則化 (Fused Lasso) [16] は、あるセンサーが観測するデータは前後する時刻で同一のクラスターに

*連絡先: NTT コミュニケーション科学基礎研究所
〒 619-0037, 京都府相楽郡精華町光台 2-4
E-mail: koh.takeuchi@lab.ne.jp

所属する (変化点ではない) という仮定から変化点を検出するために使用されている．結合正則化は前後する潜在平均値が同一の値となるよう，潜在平均値の差分に ℓ_1 -ノルムを正則化項として課す．

$$\Omega_{\text{FL}}(\mathbf{Z}) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2-1} w_{i_1, i_2} \|z_{i_1, i_2} - z_{i_1, i_2+1}\|_1, \quad (1)$$

このとき w_{i_1, i_2} は各 ℓ_1 ノルムに対する重みである．次に損失関数を定める．

$$g(\mathbf{Z}) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \|x_{i_1, i_2} - z_{i_1, i_2}\|_2^2. \quad (2)$$

結合正則化による変化点検出は，上述の損失関数と正則化項の和を最小解を求める問題となる．

$$\arg \min_{\mathbf{Z}} g(\mathbf{Z}) + \Omega_{\text{FL}}(\mathbf{Z}). \quad (3)$$

変化点は得られた \mathbf{Z} において， $z_{i_1, i_2} \neq z_{i_1, i_2+1}$ となる箇所として検出される．結合正則化は一変量時系列データの変化時刻点検出には適応可能だが，時空間変化パターンの解析には適応できない [11]．

3 提案法

3.1 高階結合正則化による時空間変化パターン検出

あるセンサー対は 24 時間毎に同一の値を取る，近くに設置されたセンサー対は同一の条件下で近い値を取る，ある現象が時間遅れで異なるセンサー対から観測されるなど，二項関係で表現できる事前知識を時空間変化パターン検出に導入したい場合を考える．しかし，結合正則化は，このような二項関係を用いることが出来ない．そこで任意の二項関係を扱えるよう結合正則化を一般化した手法である，一般化結合正則化 (generalized fused Lasso: GFL [2, 17]) を導入する．ノードが要素 $i \in \mathcal{D}$ ，エッジが要素対への重みに対応するグラフ行列を $\mathbf{W} \in \mathbb{R}^{d \times d}$ とする．このとき $d = I_1 I_2$ とする．一般化結合正則化の正則化項は次のように定義される．

$$\Omega_2(\mathbf{Z}) = \sum_{(i,j) \in \mathcal{E}} w_{i,j} \|z_i - z_j\|_1, \quad (4)$$

このとき $\mathcal{E} = \{(i, j) \mid i, j \in \mathcal{D}\}$ は全要素の任意の対からなる集合とする．したがって，一般化結合正則化は，グラフにおいて任意の要素対が隣接しているならば，要素対が同一の値をもつように作用する．

二項関係にかぎらず，より一般の多項関係を事前知識として導入したい場合も存在する．多項関係で表現で

きる事前知識には，特定の領域に設置されたセンサーグループの観測値は同一になる，あるいは同一のドメインに所属するコンピュータは同じように振る舞うなどがある．しかし，一般化結合正則化はグラフを用いているため，グループ構造は表現できない．そこで，ハイパーグラフを用いることで多項関係にあるパラメータが同一の値を持つという事前知識を導入できる高階結合正則化 (The higher-order fused Lasso: HOFL [15]) を導入する．あるグループに所属するパラメータを $g \subset \mathcal{D}$ とし，グループの集合を $\mathcal{G} = \{g_1, \dots, g_K\}$ とする．このとき， K はグループの総数とする． k 番目のグループに所属する各パラメータに対する重みを以下のように定める．

$$c_{k,i} = \begin{cases} c_k & \text{if } i \in g_k \\ 0 & \text{otherwise} \end{cases} \quad (i \in V),$$

このとき $c > 0$ とする．高階結合正則化の正則化項は次のように定められる．

$$\Omega_3(\mathbf{Z}) = \sum_{k=1}^K \left(\sum_{i \in \mathcal{D}} c_{k,i} \|z_i - \bar{z}_{k,i}\|_1 + \theta_k (z_{s_k} - z_{t_k}) \right) \quad (5)$$

このとき $\theta_k > c$ は， k 番目のグループにおけるパラメータの同一性を調整するハイパーパラメータであり， $\bar{z}_{k,i}$ は以下のように定める．

$$\bar{z}_{k,i} = z_{s_k} \text{ (if } i \geq s_k), z_{t_k} \text{ (if } i \leq t_k) \text{ and } z_i \text{ (otherwise)}$$

このときパラメータ \mathbf{Z} を降るべき順に並べ替えた際の $\mathbf{Z}_{j_1} \geq \dots \geq \mathbf{Z}_{j_d}$ をインデックスを $j_1, j_2, \dots, j_d \in V$ とし， $\theta_{\max}^k \geq \theta_0^k$ ， $\theta_{\max}^k \geq \theta_1^k$ は， $S \subset V$ ， $(\theta_0^k + c_0^k(V \setminus S) \geq \theta_{\max}^k) \vee (\theta_1^k + c_1^k(S) \geq \theta_{\max}^k) = 1$ をみたすとする．また閾値パラメータ s_k と t_k は以下のように定められるとする．

$$s_k = \min \{j' \mid \sum_{i \in \{j_1, \dots, j'\}} c_{k,i} \geq \theta\} \quad \text{and} \\ t_k = \min \{j' \mid \sum_{i \in \{j', \dots, j_d\}} c_{k,i} < \theta\}.$$

式 (5) の第 1 項は，あるグループに所属するパラメータの値が $z_{j_{s_k}}$ よりも大きい，(あるいは $z_{j_{t_k}}$ よりも小さい) パラメータが $z_{j_{s_k}}$ (あるいは $z_{j_{t_k}}$) と同一の値を持つように作用する．第 2 項は値が $z_{j_{s_k}}$ よりも大きく， $z_{j_{t_k}}$ よりも小さいパラメータがすべて同じ値を持つように作用する．このため，高階結合正則化は $(z_{j_{s_k}}, \dots, z_{j_{t_k}})$ の範囲外にあるパラメータがそれぞれ異なるの値を取ることを許容するため，グループの内，問題に適した部分のみを利用し，それ以外の部分を無視できることから，頑強なモデルとなっている．

3.2 ハイパーグラフ結合正則化

二項関係と多項関係で表現できる事前知識を共に用いた時空間変化パターン検出を行うために、ハイパーグラフ結合正則化 (Hyper-graph Fused Lasso: HFL) を提案する。HFL の正則化項を次のように定める。

$$\Omega(Z) = \lambda_1 \Omega_1(Z) + \lambda_2 \Omega_2(Z) + \lambda_3 \Omega_3(Z), \quad (6)$$

このとき、パラメータの推定値を疎にするため、Lasso の正則化項である、 $\Omega_1 = \sum_{i \in \mathcal{D}} \|z_i\|_1$ を導入し、 $\lambda_1, \lambda_2, \lambda_3$ を正則化項の重み係数とする。定義から、HFL は Lasso、結合正則化、高階結合正則化を特殊例として含む。HFL は次の問題の最小解を求める問題とする。

$$\arg \min_Z g(Z) + \Omega(Z). \quad (7)$$

HFL の目的関数は、凸関数の和からなる凸関数である。しかし、この関数は微分不可能な項を持つため最小化が困難である。また、各凸関数の最小化法が得られたとしても、これらの和の最小化は自明ではない [3]。この課題を解決するために、本研究では離散最適化を用いた最小化法を提案する。

集合 $V = \{1, 2, \dots, d\}$ に対する集合関数を $f: 2^V \rightarrow \mathbb{R}$ とする。劣モジュラ関数 [4] は次式を満たす集合関数である。

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T), \forall S, T \subseteq V. \quad (8)$$

劣モジュラ関数は最適化に関する性質を持つ。さらに劣モジュラ関数のロバシユ拡張は凸関数となる [13]。集合関数 f のロバシユ拡張 $\hat{f}: \mathbb{R}^d \rightarrow \mathbb{R}$ を

$$\hat{f}(z) = \sum_{i=1}^d z_{j_i} (f(\{j_1 \dots j_i\}) - f(\{j_1 \dots j_{i-1}\})), \quad (9)$$

とする。このとき $j_1, j_2, \dots, j_d \in V$ は、 z の要素を降るべき順に並べ替えるインデックスとする。また、劣モジュラ関数とモジュラ関数の和は、劣モジュラである [6]。なお、劣モジュラ関数 f で $f(\emptyset) = 0$ をみたすものに対して、基多面体は、

$$B(f) = \{x \in \mathbb{R}^d \mid x(S) \leq f(S) (\forall S \subseteq V), x(V) = f(V)\}. \quad (10)$$

として定められる。次にカット関数を

$$f_1(S) = \sum_{(i,j) \in E, i \in S, j \in V \setminus S} w_{i,j}, \quad (11)$$

とする。このとき E をエッジの集合、 $w_{i,j}$ を i 番目と j 番目のノード間のエッジの容量とする。一般化結合正則化の正則化項は、カット関数のロバシユ拡張と一致し [1, 17]、カット関数は二次の劣モジュラ関数とおお

よそ一致する [8]。他方、多項関係を用いた集合関数も多数存在し、Kohli らは頑強 P^n ポッツ模型を提案している。[12]。頑強 P^n ポッツ模型は次式で定められる。

$$f_2(S) = \sum_{k=1}^K \min(c_k(V \setminus S), c_k(S), \theta_k), \quad (12)$$

このとき c_k は $c_{k,i}$ を結合したベクトルとする。高階結合正則化の正則化項は、式 (12) のロバシユ拡張と一致する [15]。

本研究では、カット関数と頑強 P^n ポッツ模型の和からなる劣モジュラ関数を次のように定める。

$$f(S) = f_1(S) + f_2(S). \quad (13)$$

Lemma 1. $\Omega_2(z) + \Omega_3(z)$ は式 (13) の劣モジュラ関数のロバシユ拡張と一致する。

Proof. 式 (4) はカット関数のロバシユ拡張と一致し (11)、式 (5) は頑強 P^n ポッツ模型のロバシユ拡張と一致する [15]。ロバシユ拡張の定義 (9) より、これらの和を取ると $\Omega_2(z) + \Omega_3(z)$ が得られる。□

3.3 近接法を用いた最小化法

定義から、ハイパーグラフ正則化の正則化項は微分不可能な下半連続凸関数である。本節では、近接法 [14] を用いた最小化法を提案する。ここでは Z を並べ替えたベクトルを z とする。近接法は次のように定め、

$$\text{prox}_{\gamma\Omega} \hat{z} = \arg \min_{z \in \mathbb{R}^d} \Omega(z) + \frac{1}{2\gamma} \|\hat{z} - z\|_2^2, \quad (14)$$

近接法 $\text{prox}_{\gamma\Omega} \hat{z}$ の解を z^* とする。 ℓ_1 正則化の解は他の凸関数の解に近接法を用いることで得られる: $\text{prox}_{\Omega_1 + \Omega} = \text{prox}_{\Omega_1} \circ \text{prox}_{\Omega}$ [5]。式 $\Omega(z) = \max_{s \in B(f)} z^T s$ [6] を式 (14) に代入すると、近接法は次の基多面体上の最小化問題となる [17]。

$$\min_{z \in \mathbb{R}^d} \Omega(z) + \frac{1}{2\gamma} \|\hat{z} - z\|_2^2 \leftrightarrow \min_{s \in B(f)} \|s - \gamma^{-1} \hat{z}\|_2^2. \quad (15)$$

$t = s - \gamma^{-1} \hat{z}$ とし、劣モジュラ関数の基多面体の性質から、近接法は次の最小点問題と一致する。

$$\min_{s \in B(f)} \|s - \gamma^{-1} \hat{z}\|_2^2 = \min_{t \in B(f - \gamma^{-1} \hat{z})} \|t\|_2^2. \quad (16)$$

式変形から $z^* = -\gamma t^*$ が式 (16) の解となる。

一般に式 (16) の問題は最小ノルム点アルゴリズム [7] などの最小化法によって解が得られるが、劣モジュラ関数一般に適応可能な最小化法は最悪で $O(d^5 EO + d^6)$ (EO は関数の評価コストとする) の計算コストが

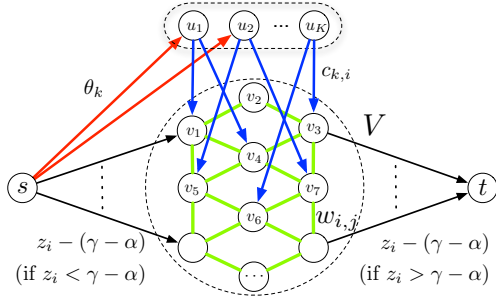


図 1: $G = (V, E)$ を HFL の正則化項に対応するグラフとする．エッジの容量は次のとおりとする． $c(s, u_k) = \theta_k$, $c(v_i, v_j) = w_{i,j}$, $c(u_k, v_i) = c_{k,i}$, $c(s, v_i) = z_i - (\gamma - \alpha)$ if $z_i > \gamma - \alpha$, $c(v_i, t) = (\gamma - \alpha) - z_i$ if $z_i < \gamma - \alpha$. u_k $k = (1, \dots, K)$ はグループに対応するハイパーノードである．ノード s, t, v_i は, ソース, シンク, パラメータの要素に対応する．

掛かるため, データ次元が大きい場合に実行不可能となる．そこで本研究では, 式 (13) がグラフ表現可能な劣モジュラ関数であることを利用し, 式 (16) の問題を図 1 の s/t 最小カット問題として最大流アルゴリズムで解く [9, 17]. . 最大流アルゴリズムは, 最悪でも $O(d|E| \log(d^2/|E|))$ の計算コストで解が得られるため大規模なデータでも実行可能である (この時, $|E|$ は図 1 のグラフのエッジ数とする).

3.4 行列補完

時空間データは欠損値を含む場合がある．マスク行列 $M \in \{0, 1\}^{I_1 \times I_2}$ によって, X の要素が欠損しているかを次のように示す．

$$m_{i_1, i_2} = \begin{cases} 1 & \text{if } x_{i_1, i_2} \text{ is observed,} \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

欠損値を考慮した損失関数を次のように定める．

$$g(\mathbf{Z}) = \sum_{i \in \mathcal{D}} m_i \|x_i - z_i\|_2^2. \quad (18)$$

損失関数とハイパーグラフ正則化項から, 時空間変化パターン検出問題を次のように定める．

$$\arg \min_{\mathbf{Z}} g(\mathbf{Z}) + \Omega(\mathbf{Z}). \quad (19)$$

定義から, この問題は 2 つの凸関数の和の最小化問題となっている．損失関数 g はリプシッツ連続凸関数であり, 正則化項は非リプシッツ連続凸関数で近接法が式 (14) によって与えられた．以上から前方-後方分離型近接点法 [3] を適応し最小解を求める．さらにネステロフ加速法を用いることで $O(1/t^2)$ の収束レートが達成する [3].

4 実験

4.1 人工データを用いた欠損値推定実験

行と列の長さが (50×50) の真値行列を作成し, 行列の要素を 12 個の領域に分割する．9 つの長方形と星形の領域を既知のグループ, 円と菱形を未知のグループとする．各領域に属する要素は同一の値を持つとし, 値は $\{0, \dots, 12\}$ のいずれかの値とした．観測行列 X は真値行列に平均 0, 分散 1 のガウス分布から生成させたノイズを加えたものとする．図 2a に真値行列と観測行列を示す．行列の要素をランダムに欠損させた．欠損要素の割合は $p = \{0.1, 0.5, 0.9\}$ とした．

比較手法には, 一般化結合正則化 (GFL), 高階結合正則化 (HOFL), 観測値の平均値 (Mean) を用いる．データに関する事前知識として, 要素の対 (i, j) が行列内で隣接していれば $w_{i,j} = 1$, そうでなければ $w_{i,j} = 0$ とする二項関係, 既知の領域に所属する要素のグループを表現した多項関係を利用する．すなわち, 要素 i が k 番目の領域に所属していれば $c_{k,i} = 1$, そうでなければ $c_{k,i} = 0$ とした．HOFL の θ_k は交差検定によって $\{1, 10, 100, 1000\}$ から選択した．提案法の重み係数 $\lambda_1, \lambda_2, \lambda_3$ は, $\{0, 0.1, 1.0, 5.0, 10.0\}$ から 10 交差検定によって選択した．

欠損値の推定精度を比較するために, Adjusted Rand index (ARI) と Root-Mean Squared Error (RMSE) を用いた．ARI はクラスタリングの精度を比較する指標で, 今回は同じ平均値を持つ要素を同一のクラスタとしてみなすことで ARI を算出した．RMSE は, 欠損値の推定精度として広く利用される指標である． $RMSE = \sqrt{\frac{\sum_{n \in \mathcal{M}} (x_n - z_n)^2}{|\mathcal{M}|}}$. このとき, \mathcal{M} は欠損値の集合とする．各設定で 10 回の試行を行い各指標の平均値を求めた．

実験の結果を表 1, 2, 3 に示す．提案法が既存法よりも良い, あるいは同等の精度を示すことを確認した． $p=0.9$ で GFL は RMSE, $p=0.1$ で GFL が ARI で提案法と同等の性能を示した．図 2 に, 真の平均値, 観測行列, 推定された潜在平均値行列を示す．提案法はグラフとハイパーグラフを利用することで, 既存法よりも精度の高い推定結果を得たことが確認された．

4.2 気象データ実験

提案手法の性能を気象データを用いた実験によって検証する．データは United States historical climatology network (USHCN) ¹ より取得した．月毎の最高平均気温 (TMAX), 平均最低気温 (TMIN), 平均気温 (TAVG), 総降水量 (PRCP) の 4 種類の観測値を利用す

¹<http://cdiac.ornl.gov/epubs/ndp/ushcn/>

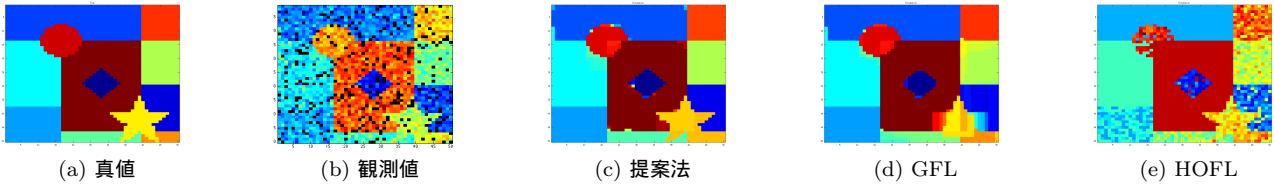


図 2: 人工データでの欠損値推定結果

表 1: 人工データ実験 ($p = 0.1$)

	Prop.	GFL	HOFL	Mean
ARI	0.94	0.90	0.88	(<0.01)
RMSE	0.32	0.32	0.65	1.54

表 2: 人工データ実験 ($p = 0.5$)

	Prop.	GFL	HOFL	Mean
ARI	0.88	0.64	0.76	(<0.01)
RMSE	0.90	0.94	3.15	2.82

表 3: 人工データ実験 ($p = 0.9$)

	Prop.	GFL	HOFL	Mean
ARI	0.88	0.23	0.88	(<0.01)
RMSE	2.45	3.45	2.49	3.74

表 4: USHCN の RMSE

	Prop.	GFL	HOFL	FL	Mean
TAVG	167	167	588	582	824
TMAX	187	188	705	702	997
TMIN	172	172	317	435	621
PRCP	278	278	333	427	637

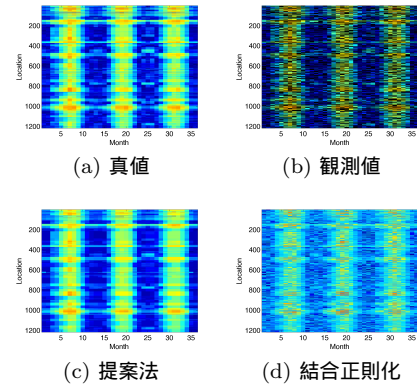


図 3: TAVG の欠損値推定結果

る。観測位置はアメリカ合衆国内の 1,218 箇所、2011 年 1 月から 2014 年 12 月までの 36 ヶ月分のデータを利用した。それぞれの観測値から $X \in \mathbb{R}^{1218 \times 36}$ の行列を作成した。時間に関する事前知識として、前後する月をグラフ、年ごとの周期性をハイパーグラフで表現した。空間に関する次元知識として、観測位置ごとの k -近傍をグラフ、週ごとのグループをハイパーグラフで表現した。データの欠損率は $p = 0.5$ とし、気象データには真のクラスタが存在しないため、RMSE で欠損値の推定精度を比較した。時間に関する構造情報のみを用いた場合の結果と比較するために比較手法に結合正則化を採用した。

実験結果を表 4 に示す。提案法と GFL が他の手法よりも優れた結果を示した。提案法の重み係数は $\lambda_2 = 10, \lambda_3 = 0$ となったため、本実験ではグラフ構造が有効であったことが分かる。図 3 に推定結果を示す。提案手法は、欠損値の推定を行いつつ、変化パターンの検出を行ったことがわかる。一方、FL は時間的な連続性のみを利用するため推定結果が乱雑であり、また空間的な変化パターンを捉えられていない。ニューヨークとフロリダに設置されたセンサーの観測値と推定値を図 4、2014 年の 8,9,10 月の推定結果を図 5 に示す。円の色は推定値を表す。図から、時空間変化パターンの検出が確認された。例えば、2014 年 9 月にニューヨークで変化が起きているが、これはフロリダでは起きている。また、図 5a, ?? から、ニューヨークで起きた変化はアメリカの東海岸から中央部にかけて広い範囲で起きたものと分かる。図 5c から、フロリダを含む東南部で変化が起き、複数のパターンが 1 つのパターンとなっていることが分かる。

5 むすび

多変量時系列データからの時空間変化パターン検出法として、ハイパーグラフ構造正則化を用いた行列補間法とその最小化法を提案した。人工データと実気象データを用いた実験から提案法の優位性を確認した。

謝辞 本研究は JSPS 科研費 14435225,14500801 の助成を受けたものです。

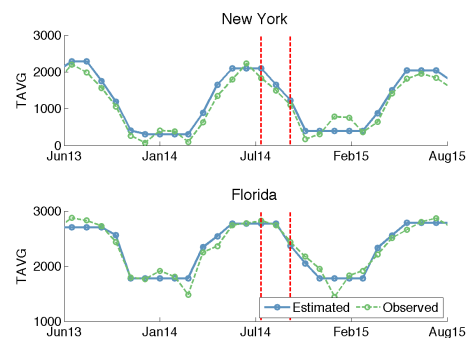


図 4: ニューヨークとフロリダの TAVG 推定結果

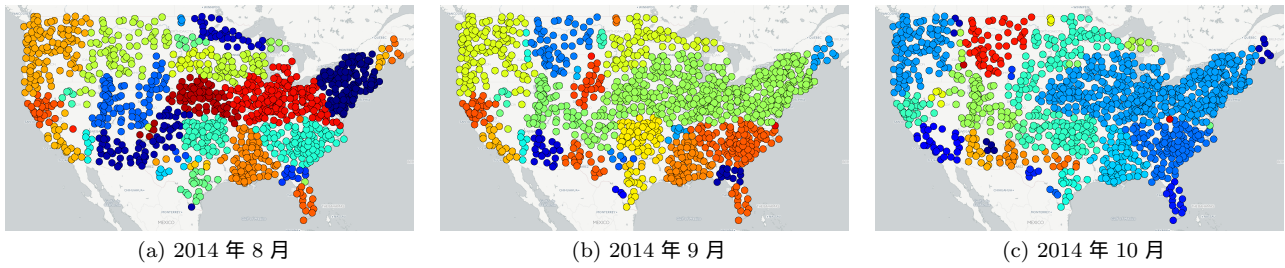


図 5: TAVG 推定結果の空間的可視化.

参考文献

- [1] F. R. Bach. Shaping level sets with submodular functions. In *Proc. of NIPS*, pages 10–18, 2011.
- [2] A. Barbero and S. Sra. Fast newton-type methods for total variation regularization. In *Proc. of ICML*, pages 313–320, 2011.
- [3] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- [4] J. Edmonds. Submodular functions, matroids, and certain polyhedra. *Combinatorial Structures and Their Applications*, pages 69–87, 1970.
- [5] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani, et al. Pathwise coordinate optimization. *The Annals of Applied Statistics*, 1(2):302–332, 2007.
- [6] S. Fujishige. *Submodular functions and optimization*, volume 58. Elsevier, 2005.
- [7] S. Fujishige, T. Hayashi, and S. Isotani. The minimum-norm-point algorithm applied to submodular function minimization and linear programming. Technical report, Research Institute for Mathematical Sciences Preprint RIMS-1571, Kyoto University, Kyoto, Japan, 2006.
- [8] S. Fujishige and S. B. Patkar. Realization of set functions as cut functions of graphs and hypergraphs. *Discrete Mathematics*, 226(1):199–210, 2001.
- [9] G. Gallo, M. D. Grigoriadis, and R. E. Tarjan. A fast parametric maximum flow algorithm and applications. *SIAM Journal on Computing*, 18(1):30–55, 1989.
- [10] V. Guralnik and J. Srivastava. Event detection from time series data. In *Proc. of SIGKDD*, pages 33–42, 1999.
- [11] Z. Harchaoui and C. Lévy-Leduc. Multiple change-point estimation with a total variation penalty. *Journal of the American Statistical Association*, 2012.
- [12] P. Kohli, L. Ladicky, and P. H. S. Torr. Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision*, 82(3):302–324, 2009.
- [13] L. Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer, 1983.
- [14] J. J. Moreau. Fonctions convexes duales et points proximaux dans un espace hilbertien. *CR Acad. Sci. Paris Sér. A Math*, 255:2897–2899, 1962.
- [15] K. Takeuchi, Y. Kawahara, and T. Iwata. Higher order fused regularization for supervised learning with grouped parameters. In *Proc. of ECMLPKDD*, pages 577–593, 2015.
- [16] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108, 2005.
- [17] B. Xin, Y. Kawahara, Y. Wang, and W. Gao. Efficient generalized fused lasso with its application to the diagnosis of alzheimer’s disease. In *Proc. of AAAI*, pages 2163–2169, 2014.
- [18] X. Zhou, S. Shekhar, and R. Y. Ali. Spatiotemporal change footprint pattern discovery: an interdisciplinary survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(1):1–23, 2014.

11月9日(水) 発表分の受賞理由

1. SIG-SLUD：言語・音声理解と対話処理研究会

「説明行為の質の推定に向けた会話者のマルチモーダル情報モデリング」

岡田将吾，米航，新田克己（東京工業大学）

（SIG-SLUD-B501-07，第 74 回，2015 年 7 月 22 日）

選考理由：

本研究は，人間同士のグループ会話の参加者が他の参加者に説明を行う場面において，その説明行為の質の推定を行う手法を提案し，有効性を示している。説明の質の推定という新しいタスクに焦点を当てた点に，研究として新規性があり，タスクの有用性も高い。予め収集した会話データに対して，3名の第三者が説明の質を評定した。説明の質の評定は一見主観性が高く評定者間で一致しないことが危惧されるが，同一の説明課題についてのデータを多数のグループで集めることで，十分な一致率が得られることを示している。提案手法では，話し手・聞き手が産出したアトミックな言語・非言語行為の頻度・時間長を中心とした統計情報に加えて，これらアトミックな種々の行為の間の話者内（モダリティ間）・話者間での共起頻度を用いている点に特徴があり，そのために共起クラスタリング手法を用いて，注目するパターンの数を絞る方法も示している。また，言語情報については作業者による書き起こしを用いているものの，その他の入力情報については，モーションセンサデータから自動抽出した結果を用いている。このため，観察可能な情報と説明の質の間の関係性を人文学的立場から明らかにしているだけでなく，推定技術としての実現可能性も示している。今後，音声認識・画像認識に基づく完全自動推定技術への展開が期待される。研究会での参加者の反応もよく，本タスクへの関心の高さがうかがわれ，今後，後続の研究が期待できる。以上を総合的に勘案し，本研究が受賞に値するという判断に至った。

2. SIG-Challenge：AI チャレンジ研究会

「複数のマイクロホンアレイとロボット聴覚ソフトウェアHARK を用いた野鳥の観測精度の検討」

松林 志保（名古屋大学），小島 諒介（東京工業大学），中臺 一博（東京工業大学／(株) ホンダ・リサーチ・インスティテュート・ジャパン），鈴木 霊璽（名古屋大学）

（SIG-Challenge-043-11，第 43 回，2015 年 11 月 12 日）

選考理由：

本研究は，屋外フィールドでの野鳥の行動分析を扱っている。これまで，野鳥の行動分析は，実際に人が現場で自分の耳で聞きながら，鳥の位置や種類を紙に記入することで行っていたため，手間がかかる，また，信頼性が保証できないといった問題があった。本研究では，ロボット聴覚ソフトウェアHARK を利用したマイクロホンアレイ処理を行うことにより，木立の立て込んだ複雑な屋外で野鳥の位置を検出することを試み，その有効性を示すとともに適用範囲や課題を明らかにするため，詳細にわたる実験結果の分析を行っている。複数の研究分野にまたがるチャレンジングな研究であり，研究としての信頼性，新規性も高いことから優秀賞の候補にふさわしいと考える。

3. SIG-KST 知識・技術・技能の伝承支援研究会

「人の流れの計測とシミュレーションによる避難誘導方法の伝承－新国立劇場における避難体験オペラコンサートを例に－」

大西正輝（産業技術総合研究所），山下倫央（産業技術総合研究所），星川哲也（新国立劇場），佐藤和人（新国立劇場）

（SIG-KST-026-06，第 26 回，2015 年 11 月 13 日）

選考理由：

本論文は、大規模な人数を集めての避難訓練を対象に、人の流れの計測技術とシミュレーション技術に基づいて人の流れを正確に理解し、得られた知識や知見を正しく伝承するための手法について述べている。その中で、2014 年に新国立劇場で実施された1300 人規模の避難体験オペラコンサートでの実証実験を対象に、提案手法による計測結果とシミュレーション結果、およびその分析によって得られた知見が示され、実際に大規模な避難訓練に適用して有効であることが示されている。以上のように、学術的にも価値のある研究テーマである人の流れの計測技術とシミュレーション技術に取り組み、大規模な避難訓練の知識伝承という社会的課題に対して多大な貢献が期待できるため、受賞に値する。

4. SIG-FPAI 人工知能基本問題研究会

「大規模グラフのコンパクトでスケーラブルな全距離スケッチ」

秋葉拓哉，矢野洋祐（国立情報学研究所）

（SIG-FPAI-B503-06，第 100 回，2016 年 3 月 27 日）

選考理由：

本論文は、大規模グラフ上で様々なクエリ処理に必要な基本演算を、グラフに対して事前に前処理を行うことで、メモリと速度の両面でコストをきわめて小さくすることができるデータ構造(SRS)を提案している。既存手法では、前処理にグラフサイズと同程度の領域が必要なデータ構造(ADL)を構築していたが、SRS では、このADL を経由しない軽量な適応型アルゴリズムを開発し、実際の大規模グラフにおいて領域計算量を大幅に改善できることを示している。本研究は、今後ますます大規模化するネットワーク解析に対して重要な貢献が期待される優れた研究である。

5. SIG-FPAI 人工知能基本問題研究会

「高階結合正則化による時空間変化パターン検出」

竹内孝（日本電信電話（株）），河原吉伸（大阪大学），岩田具治（日本電信電話（株））

（SIG-FPAI-B503-10，第 100 回，2016 年 3 月 27 日）

選考理由：

本論文は、多変量の時系列データに対してその変化パターンの抽出を行う手法の提案とその応用の研究である。多変量の時系列データ時間軸に関する時系列としての変化と多変量の各要素が示す位置情報などの変化と二軸的表現から次元が高くなり実用的な変化パターンの抽出法の設計は技術的な課題である。本論文では平均値が変化する変化点検出としての解釈がしやすい自然な定式化、および、二項関係および多項関係を共に構造正則化の制約として課すことで自然な結果が得られる手法を考案している。また劣モジュラ性に基づき効率的な計算アルゴリズムを導出しており、今後様々な課題への応用が期待でき、優秀賞に値する優れた研究である。

記述論理に基づく農作業オントロジーの設計と応用

A Design and Applications of Agricultural Activity Ontology based on Description Logic

朱 成敏^{1*} 小出 誠二² 武田 英明¹ 法隆 大輔³
竹崎 あかね³ 吉田 智一³

Sungmin JOO¹ Seiji KOIDE¹ Hideaki TAKEDA¹ Daisuke HORYU³
Akane TAKEZAKI³ Tomokazu YOSHIDA³

¹ 国立情報学研究所

¹ National Institute of Informatics

² 情報・システム研究機構

² Research Organization of Information and Systems

³ 国立研究開発法人農研機構中央農業総合研究センター

³ National Agriculture and Food Research Organization

Abstract: The spread of agricultural IT systems has been accompanied by generating various data such as farming information or agricultural activities. These data is expected to optimization of agricultural activities and prediction of harvest yield. To analyze the stored data, data sharing and integration between agricultural IT systems is essential. However, since there is a lot of multi-meaning vocabulary and no guidelines in agriculture field, each of the agricultural IT systems has been operated based on its own vocabulary. In this paper, we propose and we develop agricultural activity ontology based on description logic as core vocabulary for agriculture that becomes the guidelines for data federation between agricultural IT systems.

1 はじめに

近年、日本の農業分野において IT システムが普及されており、作物の生産や営農において効率の良い管理が可能となった。また、農業 IT システムから発生されたデータは収穫量の予測や営農計画において重要な資料であり、これらのデータを蓄積し、統計・分析することによって農業分野における全体的な把握や予測の指標として用いることができる。

しかし、農業 IT システムは基準となる情報が標準化されていないため、IT ベンダーはそれぞれ独自の用語を用いてデータの項目を構成している。そのため農業 IT システムから発生された農業データは統合が困難となり、全体的な統合・分析において障害となっている。現在、38 システムが農業経営、生産管理に係わっており、今後の普及のためにも農業情報の標準化が必要である [1]。特に農作業の名称は作業の計画など営農管理において基本となる項目であり、統計・分析において

も最も重要な項目である。しかし、農作業の名称は同じ作業でも環境や習慣によって異なる場合や「整地作業」と「代かき」のように作業の範囲を決めずに表記される場合などその標準化が進めていない状況である。こういった背景から内閣府 IT 総合戦略本部は農業情報の総合運用性と可能性の確保のために標準化や情報の標準化や情報の取扱いに関する政府横断的な戦略を策定し、その戦略ロードマップから「農作業の名称」を優先的に導入する課題として選定された [2]。

本稿では「農作業の名称」の標準化のために要求される特徴について考察し、その考察から発見した特徴を記述論理を用いて定義を行う。そして、本稿が提案する農作業オントロジーについて考察し、農作業名称の標準語彙体系としての可能性について述べる。

2 農作業に関する既存の語彙体系

本研究ではまず、農作業オントロジーを設計するために農作業に関する既存の語彙体系について分析を行った。対象となる既存の語彙体系は国際連合食糧農業機

*連絡先：国立情報学研究所
〒101-8430 東京都千代田区一ツ橋 2-1-2
E-mail: joo@nii.ac.jp

関(FAO)が管理するAGROVOCと内閣府が提示した「農業ITシステムで用いる農作業の名称に関する個別ガイドライン(試行版)(以下ガイドライン)」である。

AGROVOCは国際連合食糧農業機関と欧州共同体委員会(CEC)が共同で開発した農林水産,食糧,環境,およびその関連分野の専門用語を32,000語以上網羅したシソーラスであり,21言語¹で収録されている。国際的な標準ともなっている[3]。農作業に関連する概念は最上位概念である「activities」に含まれており,1,434語の下位概念が収録されており,その中で日本語の表記がある概念は1,221語である²。AGROVOCは国際的に相互利用性を持つ標準語彙体系であるが,イネ作に関する作業など日本の農業分野で重要とされる農作業に関してはその収録が不足している。例えば,稲農作業語での重要概念である「苗取り」や「中干し」,「代かき」などはAGROVOCに収録されていない。また,日本語表記が収録されていない語彙も多く存在している。何より,AGROVOCは広義語(BT)と狭義語(NT)による階層構造のため兄弟関係の語彙についてはその関係性が曖昧である。例えば「収穫」の下位概念である「草刈」と「藁むき」は作業の一部として「収穫」との関係があるが,兄弟関係として関連性があるとは言い難い。こういった基準の曖昧さは新規用語が追加される際,効率的な対応に妨げとなる。

内閣府が農作物名称の標準化のために提案した「ガイドライン」では農作業名称が大・中・小項目に分類されている[4]。大項目では「は種・定植」,「耕起整地」,「出荷」,など14項目で構成されており,中項目は「は種・覆土」,「苗取り・運搬」などより具体化して分類されている。小項目は利用者が適宜拡張を行うことが可能となり,農作業名称が持つ多様性に柔軟な対応が可能である。しかし,こういった柔軟性はITシステム間のデータ共有や統合において互換性が確保できない可能性がある。また,階層間や項目間に同一項目が存在する場合もあり,階層構造において曖昧さがある。

標準語彙体系としての既存語彙体系は概念間の関係の曖昧性,農作業名称が持つ語彙の多様性という課題を持つ。また,階層構造の曖昧さもあり,用語の明確な定義と分類に影響を与えられ考えられる。今後新規語彙が追加される場合を考慮して明確な階層構造を構築する必要がある。これらの課題を改善するために農作業オントロジーを構築し,農作業名称の標準化に用いる方法が挙げられる。概念間の関係を細分化し,また概念と用語表記を分離することで,明確に概念を定義することが可能なオントロジーは,農作業の名称が持つ特徴をより明確に反映することができる。そこで,我々は農作業における標準語彙体系のために農作業オントロジーの構築を進めてきた[5][6]。

¹2015年9月24日時点

²2015年2月19日時点

3 農作業オントロジーの設計

我々は前章で述べた農作業が持つ特徴に基き「農作業基本オントロジー(AAO:Agriculture Activity Ontology)」を構築,公開してきた³。公開版のver 0.94⁴では「目的」と「手段」,2つの属性によって概念の分類を行った[5][6]。また,ver 1.00⁵では属性を具体化し,その継承関係も明確にした基準で分類を行った。そして,記述論理に基づき,概念階層の記述を論理的に定義することで概念の矛盾や同一性を明確にした。最新版であるver 1.10⁶では既存の農業現場で行われている作業を追加・補完し,より実用性を高めた。本章では農作業基本オントロジーの特徴について詳しく述べる。

3.1 農作業の構造化

農作業の概念を構造化するために農作業を作物生産に係わる作業である「作物生産作業」,経営などを含む「営農管理作業」の2つを上位概念とした。そして,「作物生産作業」は作物の生育を目的とする「作物生育作業」,環境の制御を目的とする「環境制御作業」,収穫後農産物の生産を目的とする「収穫後作業」,作物生産において間接的な支援を目的とする「作物生産支援作業」の4つに分類される。このように,それぞれの農作業が持つ共通の「目的」を基準として概念の分類を行った。

そして,より具体的に概念を定義するために「行為」,「対象」,「場所」,「手段」,「時期」を,特定作物に関連する作業を定義するために「作物」を属性として用いた。これらの属性を用いて農作業の概念を記述し,概念間の分類を行うための基準とした。図1は繁殖に関連する作業の分類例である。

例えば「は種」という作業について記述論理を用いて定義する場合は以下ようになる。まず「は種」は「種子繁殖作業」の一部であり「種子繁殖作業」は「繁殖制御作業」を「繁殖制御作業」は「作物生育作業」を「作物生育作業」は「作物生産作業」を上位概念として持つ。そして「作物生産作業」の上位概念は最上位概念である「農作業」である。

- | | | | |
|--------|---|--------|-------------|
| 作物生産作業 | ≡ | 農作業 | |
| | | ∩∇ | 目的.作物生産 (1) |
| 作物生育作業 | ≡ | 作物生産作業 | |
| | | ∩∇ | 目的.作物生育 (2) |
| 繁殖制御作業 | ≡ | 作物生育作業 | |
| | | ∩∇ | 目的.繁殖制御 (3) |

³<http://www.cavoc.org/aao/>

⁴2015年5月12日公開,185語収録

⁵2015年11月2日公開,301語収録

⁶2016年2月12日公開,330語収録

	第一階層	二階層(概)	第三階層	第四階層	第五階層	作物	目的	行為	対象	場所	手段	時期
作物生産作業												
	作物生育作業						作物生産					
		繁殖制御作業					繁殖制御					
			繁殖準備作業				繁殖準備					
				採種				採る	種			
				苗取り				取る	苗			
			種子繁殖作業				種子繁殖					
				は種				播く	種	圃場		
					苗箱播種	イネ				苗箱		
					湛水直播	イネ				水田		
					乾田直播	イネ				乾田		
				覆土				被せる			土	
			栄養繁殖作業				栄養繁殖					
				挿し木				挿す	木			
				接ぎ木				接ぐ	木			
					呼び接ぎ	野菜			茎		活着後切断	
					割り接ぎ	野菜			茎		差し込み	
				取り木				取る	枝葉木			
				株分け				分ける	株			
					芋切り	パレイショ			種芋			
				分球				切り分ける	球根			

図 1: 繁殖制御作業に関する作業の例

種子繁殖作業 ≡ 繁殖制御作業
 $\cap \forall$ 目的.種子繁殖 (4) 従って農作業における「は種」は式(6)と式(7)により次のように定義される.

式(1)と式(2),式(3)は,

繁殖制御作業 ≡ 農作業
 $\cap \forall$ 目的.種子繁殖
 $\cap \forall$ 目的.繁殖制御
 $\cap \forall$ 目的.作物生育
 $\cap \forall$ 目的.作物生産 (5)

となる.ここで目的の値である「種子繁殖」,「繁殖制御」,「作物生育」,「作物生産」は定義により次のような包含関係を持つ.

種子繁殖 \sqsubseteq 繁殖制御 \sqsubseteq 作物生育 \sqsubseteq 作物生産 (6)

式(5)により以下の式が導かれる.

種子繁殖作業 ≡ 農作業
 $\cap \forall$ 目的.種子繁殖 (7)

一方「は種」は「種子繁殖」を目的とし,場所を「圃場」,対象を「種」,行為を「播く」とする作業なので以下のように定義される.

は種 ≡ 種子繁殖作業
 $\cap \forall$ 行為.播く
 $\cap \forall$ 対象.種
 $\cap \forall$ 場所.圃場 (8)

は種 ≡ 農作業
 $\cap \forall$ 目的.種子繁殖
 $\cap \forall$ 行為.播く
 $\cap \forall$ 対象.種
 $\cap \forall$ 場所.圃場 (9)

そして「は種」の中で「イネ」作において行われる作業であり,場所が「水田」の場合は「湛水直播」,「乾田」の場合は「乾田直播」として分類される.従って「湛水直播」,「乾田直播」は兄弟関係であり,「は種」の下位概念となる.ここで「湛水直播」と「乾田直播」を記述論理を用いて表現すると以下ようになる.

湛水直播 ≡ は種
 $\cap \forall$ 作物.イネ
 $\cap \forall$ 場所.水田 (10)

乾田直播 ≡ は種
 $\cap \forall$ 作物.イネ
 $\cap \forall$ 場所.乾田 (11)

場所の値である「水田」と「乾田」は「圃場」の一部である.

水田 \sqsubseteq 圃場, 乾田 \sqsubseteq 圃場 (12)

式(11)より式(12),式(13)は次のようになり「湛水直播」と「乾田直播」という農作業の定義となる.

湛水直播 ≡ 農作業
 □目的.種子繁殖
 □行為.播く
 □対象.種
 □作物.イネ
 □場所.水田 (13)

□行為.砕く
 □場所.水田)
 □ (農作業
 □目的.土壌整地
 □行為.平らにする
 □場所.圃場)
 □ (農作業 □目的.保水) (15)

そして、式 (15) の選言形式から連言形式への変換により次のことがわかる。

乾田直播 ≡ 農作業
 □目的.種子繁殖
 □行為.播く
 □対象.種
 □作物.イネ
 □場所.乾田 (14)

代かき ≡ 農作業
 □目的.(土壌整地 □ 保水)
 □行為.(砕く □ 平らにする)
 □場所.水田 (16)

このように農作業基本オントロジーでは概念が持つ属性を細分化することで階層構造を定義し、概念間の関係性を構造化した。

このように農作業基本オントロジーでは複数の目的で行われる作業のために概念間の多義的概念を考慮した。

3.2 多義的概念

農作業の中では複数の目的によって行われる作業が多数存在している。例えば、土の表面を有機物などで覆う作業である「マルチング」は温度を適度に保持・調整する同時に雑草を抑制する効果も期待される。そして、水田を砕き作土中の土塊を粉碎する作業である「代かき」は、漏水を防止するので「保水」のために、また土壌を膨軟にするので「均平化」のためにも行われる(図2)。ここで「代かき」が継承した上位概念によって表現すると以下ようになる。

3.3 別名の表記

農作業の名称は同じ作業でも様々な別名が存在する場合がある。別名が存在する場合は大きく3つに分かれる。まず、作業名に含まれる漢字が常用漢字ではない場合、例えば「播種」はそれぞれ「は種」として表記される場合がある。

は種 ≡ 播種
 ≡ 種まき (17)

そして、英語の作業名と日本語の作業名が混在する場合がある。環境順化を目的とし、苗を外気に晒す作業である「ハードニング」は英語の表記以外にも「硬化」とも呼ばれる。

ハードニング ≡ 硬化 (18)

最後に同じ作業でありながら判断の基準によって名称が異なる場合もある。例えば、秋に行われる耕起である「秋耕」は収穫後の土を深く掘り起こすため「荒起こし」とも呼ばれる。

秋耕 ≡ 荒起こし (19)

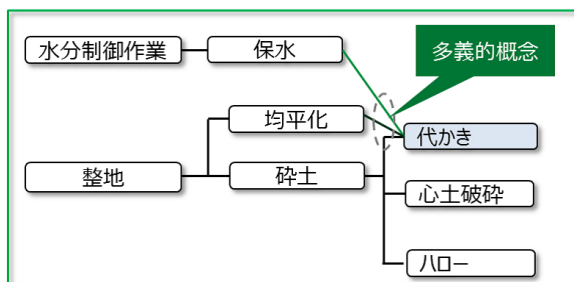


図 2: 「代かき」における多義的概念

代かき ≡ (農作業
 □目的.土壌整地

このように農業分野における用語の多様性に対応するために農作業オントロジーでは見出し語となる「概念」と別名を表す「表記」の両方を定義する必要がある。

4 農作業オントロジーの応用

本章では前章で述べた設計に基づいて構築した農作業オントロジーの応用について述べる。

4.1 URI

農作業基本オントロジーは固有の名前空間を持ち、各々農作業名はその上で定義される。図3は「は種」が持つ名前空間の例である⁷。名前空間では上位概念と下位概念と共に、作業名との関係が分かるように階層構造を表示している。また、別名の表記とその概念が持つ属性を表示することで作業の詳細が把握できるように実装された。



図 3: 農作業基本オントロジの概念が持つ名前空間

4.2 語彙リストの公開

農作業オントロジーの目的は農業分野における農作業名称の標準化である。そこで農作業基本オントロジー ver1.10 を用いて語彙リストを自動作成し、Web 上で公開している⁸。語彙リストは 3 種類があり、作業を網羅した用語リスト、概念を網羅した概念リスト、そして用語や概念の属性を定義した属性リストがある。ここで用語とは、「代かき」や「プラウ耕」など農業現場で行われる実際の作業の名称を示す。これらの用語を分類するために抽象的な概念が必要となる。例えば、用語である「代かき」と「プラウ耕」はそれぞれ「整地」、
「耕起」という上位概念に属しており、またその上の階

⁷<http://www.cavoc.org/aao/ns/1/は種>

⁸<http://www.cavoc.org>

層では「土壌整地作業」という上位概念を共通で持つ。「土壌整地作業」は属性として目的のみ持つ抽象的な概念である。このように用語の上位概念として目的を基準とする概念を本研究では概念と呼ぶ。属性とは、目的、行為、対象など用語と概念を記述する値を示す。属性は分類の基準であり、用語や概念を定義する値として用いられる。図4は属性を用いて自動生成された用語リストと概念リストの例である。用語である「緑肥鋤込み」は目的を「有機物施用」、手段を「緑肥」、行為を「混ぜる」とする作業として定義されたため用語リストでは「有機物施用作業の一つで緑肥を用いて混ぜる作業」と自動生成される。そして概念である「化学的有害動物抑制作業」は上位概念が「有害動物抑制作業」であり、目的を「有害動物制御」、手段を「化学的手段」とするため「有害動物制御を目的とし、化学的手段を手段とする有害動物抑制作業」と自動生成された。

このように農作業オントロジーが持つ概念と属性を用いることで用語の明確な定義が可能となり、農作業名称の標準化に用いることができる。

4.3 農作業語彙を支えるオントロジー推論

一般に、抽象的な概念や複合概念についてその意味的な構造を考えることは難しいが、具体的な農作業であればあるほど、その属性を列挙することは容易である。そしてオントロジーに個々の単一属性についてしっかりとした階層構造がある場合には、複数の属性の連合である概念を推論によって導くことができる。

例えば、「かかし作り」の目的は目的語彙中の「有害動物抑制」であることが判明すれば、次のように「かかし作り」を定義することができる。ここで、「かかし作り」の上位概念は定義されていないことに注意されたい。

$$\begin{aligned} \text{かかし作り}' &\equiv \vee \text{目的.有害動物抑制} \\ &\quad \cap \vee \text{行為.作る} \\ &\quad \cap \vee \text{対象.かかし} \\ &\quad \cap \vee \text{手段.物理的手段} \quad (20) \end{aligned}$$

一方に属性のオントロジーとして次のような階層があり、

$$\text{有害動物抑制} \sqsubseteq \text{有害動物制御} \sqsubseteq \text{生物制御} \quad (21)$$

他方に農作業オントロジーとして以下に示す階層構造があるため、

$$\begin{aligned} \text{物理的有害動物抑制作業} &\equiv \text{有害動物抑制作業} \\ &\quad \cap \vee \text{手段.物理的手段} \quad (22) \end{aligned}$$

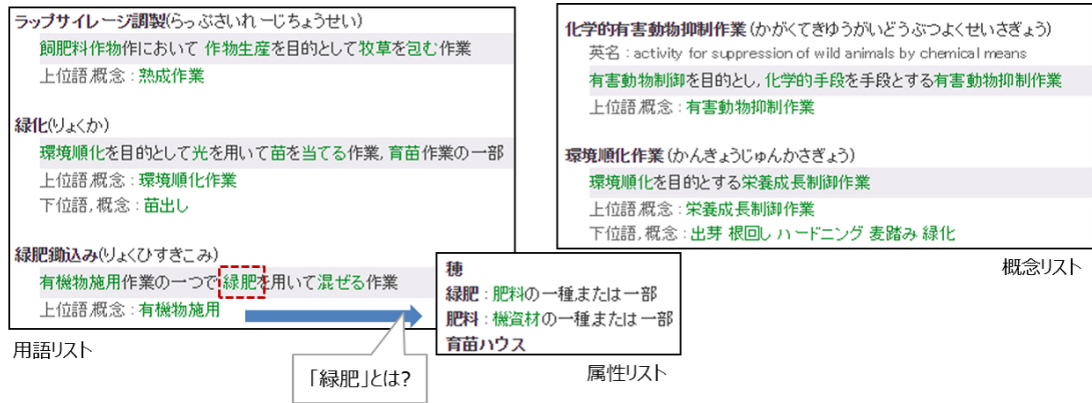


図 4: 農作業基本オントロジに基づく語彙リストの生成

有害動物抑制作業 ≡ 生物制御作業
 □∇ 目的.有害動物抑制 (23)

生物制御作業 ≡ 環境制御作業
 □∇ 目的.生物制御 (24)

先に定義した「かかし作り」の直接上位概念は「有害動物抑制(目的)」と「物理的手段(手段)」の連合である「物理的有害動物抑制作業」として矛盾がなく、問題がない。

新しい用語を農作業語彙に追加するにあたって、既存の農作業オントロジーに基づいて新しい用語の定義を支援するための簡単な推論エンジンを Schank らの Memory Organization Package(以下 MOP)[7][8] を参考に開発した。ただし、MOP[7] では概念(クラス)と個物(インスタンス)の間に区別がないが、MOP[8] では区別する。述語論理においても概念と個物は区別される。具体的な知識記述は述語論理に従う OWL[9](具体的記述構文は Turtle[10]) によって行った。式(10)の Turtle 表記を以下に示す。

```

cavoc.aao:乾田直播 a rdfs:Class ;
  rdfs:subClassOf cavoc.aao:は種 ,
  [ a owl:Restriction ;
    owl:onProperty cavoc:作物類 ;
    owl:allValuesFrom cavoc:イネ ] ,
  [ a owl:Restriction ;
    owl:onProperty cavoc:場所 ;
    owl:allValuesFrom cavoc:乾田 ] .
    
```

我々は OWL の推論システムとして、既に Common Lisp Object System (CLOS) を OWL 推論用に改造した SWCLOS[11] を所有している。その詳細は省くが、今回 SWCLOS 上に OWL 意味論に合致した MOP 機能を実装した。式(20)の定義から MOP の機能によってその直接上位概念(クラス)を導く推論の実施例を以下に示す。

```

gx(8): (get-form cavoc.aao:かかし作り)
(owl:Class cavoc.aao:かかし作り
 (rdfs:subClassOf
  (owl:Restriction _:g1937
   (owl:onProperty cavoc:目的)
   (owl:allValuesFrom cavoc:有害動物抑制))
  (owl:Restriction _:g1938
   (owl:onProperty cavoc:行為)
   (owl:allValuesFrom cavoc:作る))
  (owl:Restriction _:g1939
   (owl:onProperty cavoc:対象)
   (owl:allValuesFrom cavoc:かかし))
  (owl:Restriction _:g1940
   (owl:onProperty cavoc:手段)
   (owl:allValuesFrom cavoc:物理的手段)))
 (rdfs:label "\"かかし作り\"@ja")
 (cavoc:yomi "\"かかしつくり\"@ja"))
    
```

```

gx(9): (refine-abstraction-from
  cavoc.aao:作物生産作業 cavoc.aao:かかし作り)
#<node cavoc.aao:物理的有害動物抑制作業>
gx(10): (get-form cavoc.aao:かかし作り)
(owl:Class cavoc.aao:かかし作り
 (rdfs:subClassOf cavoc.aao:物理的有害動物抑制作業
  (owl:Restriction _:g1937
   (owl:onProperty cavoc:目的)
   (owl:allValuesFrom cavoc:有害動物抑制))
  (owl:Restriction _:g1938
   (owl:onProperty cavoc:行為)
   (owl:allValuesFrom cavoc:作る))
  (owl:Restriction _:g1939
   (owl:onProperty cavoc:対象)
   (owl:allValuesFrom cavoc:かかし))
  (owl:Restriction _:g1940
   (owl:onProperty cavoc:手段)
   (owl:allValuesFrom cavoc:物理的手段)))
 (rdfs:label "\"かかし作り\"@ja")
 (cavoc:yomi "\"かかしつくり\"@ja"))
    
```

gx(8) の get-form は式(20)に相当する計算機表現を表示させている。gx(9) の refine-abstraction-form では「cavoc.aao:作物生産作業」を出発点として概念「cavoc.aao:かかし作り」の上位概念として最も特殊な概念が探索され、その答えとして「cavoc.aao:物理的有害動物抑制作業」が得られた。gx(10) では推論結果として「cavoc.aao:かかし作り」の上位概念が追加

されたものが表示された。すなわち以下の結果が推論された。

かかし作り' ⊆ { (∨ 目的.有害動物抑制
 ∩∨ 行為.作る
 ∩∨ 対象.かかし
 ∩∨ 手段.物理的手段)
 ∩ 物理的有害動物抑制作業 } (25)

この推論の結果から「かかし作り」は以下のように定義される。

かかし作り ≡ { (∨ 目的.有害動物抑制
 ∩∨ 行為.作る
 ∩∨ 対象.かかし
 ∩∨ 手段.物理的手段)
 ∩ 物理的有害動物抑制作業 } (26)

実際「かかし作り」は2012年度の農業経営統計調査[12]で米生産の作業分類に収録されており、農作業基本オントロジーではver 1.10に新しく収録された作業名である(図5)。このように農作業基本オントロジーを用いることで適切な分類だと確認することができ、効率的な新規用語の追加が可能となる。

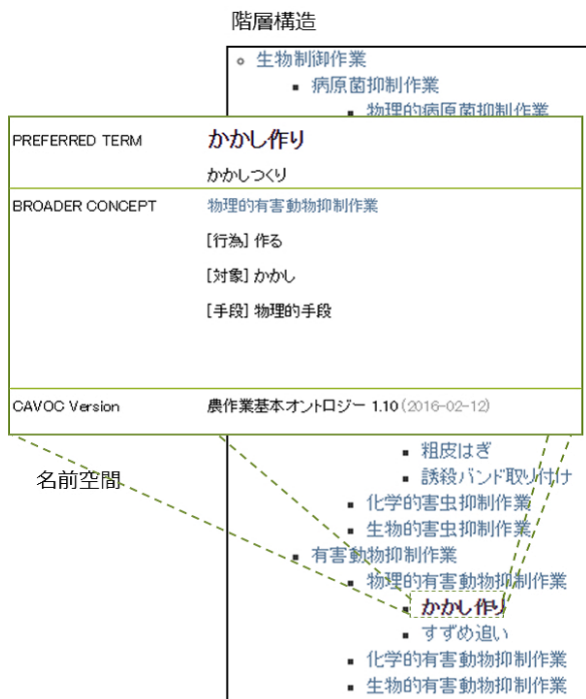


図 5: 「かかし作り」の名前空間

5 考察と今後の課題

農作業基本オントロジーの最新版である ver1.10 では概念と用語を合わせて 330 語が収録されている。語彙の収集は統計資料に基づいて行われ、また我々の調査により属性の値が定義されたため一部農業現場における検証が必要とされる部分がある。現在、その検証作業を一部進めており、今後合意を通じて確立していきたい。

農作業基本オントロジーは作物に特化した作業を含め農作業全般における作業を網羅した。作物に特化した作業は現在イネ、バレイショ、メロンなど 10 種類の作物または作物類における作業が収録されている。作物類を属性として用いるため特定作物に特化した農作業オントロジーへの拡張が可能となる。多様な作物に対応するためには作物間の関係性を定義、分類する作物基盤のオントロジーが必要となる。現在、我々は「農作物オントロジー」の構築も進めている。

6 むすび

本稿では農作業名称の標準化のために農作業オントロジーの構築を行った。オントロジーを用いることで農作業の名称が持つ多様性に明確な定義をすることができ、農作業名称の標準としての可能性を確認した。今後農作物や流通、農薬などの農業分野の用語標準化にオントロジーを積極的に取組む予定である。

謝辞

本研究の一部は委託試験研究「データ・機能のオープン化と連携による多圃場営農管理システムの開発」による研究成果に基づく。

参考文献

- [1] 内閣府, 農業情報創成・流通促進戦略に係る標準化ロードマップ(案), <http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou/dai10/siryou1-2.pdf> 2015年9月21日参照
- [2] 内閣府, 農業情報創成・流通促進戦略, <http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/senryakuzenbun_140603.pdf> 2015年9月21日参照

- [3] AGROVOC Multilingual agricultural thesaurus , <<http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>> 2015 年 9 月 21 日参照
- [4] 内閣府, 農業 IT システムで用いる農作業の名称に関する個別ガイドライン (試行版) (案), <http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou/dai10/siryou1-3.pdf> 2015 年 9 月 21 日参照
- [5] 竹崎あかね, 法隆大輔, 武田英明, 朱成敏, 吉田智一, 情報共有に資する稲作業語彙体系の試作, 農業情報学会 2015 年度年次大会, 個-14, 2015.
- [6] 朱成敏, 武田英明, 法隆大輔, 竹崎あかね, 吉田智一, 農業 IT システム間データ連携のための農作業基本オントロジーの構築, 電子情報通信学会知的環境とセンサネットワーク研究会, ASN2015-75, 2015.
- [7] Riesbeck, Christopher K., Roger C. Schank, *Inside Case-Based Reasoning*, ISBN 0-89856-767-6, LEA (1989)
- [8] Schank, Roger C., Alex Kass, Christopher K. Riesbeck, *Inside Case-Based Explanation*, ISBN 0-8058-1029-3, LEA (1994)
- [9] OWL Web Ontology Language Guide, W3C Recommendation 10 February 2004, <https://www.w3.org/TR/owl-guide/>
- [10] RDF 1.1 Turtle, Terse RDF Triple Language, W3C Recommendation 25 February 2014, <https://www.w3.org/TR/turtle/>
- [11] Seiji Koide, Theory and Implementation of Object Oriented Semantic Web Language, NII 論文 ID:500000547373, 2011
- [12] 農業経営統計調査/農産物生産費/確報/平成 24 年産米及び麦類の生産費/年次/2012 年/「利用者のために」, 農林水産省, 政府統計の総合窓口, <<http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001117795>> 2016 年 1 月 29 日参照

情報アクセスにおける受動性と能動性: 音声対話によるニュース記事アクセス

Intentionality in Information Access Behavior: A Spoken Dialogue System for Interactive Access to News Articles

林 良彦^{1*} 藤江 真也^{2,1} 福岡 維新¹ 高津 弘明¹ 小林 哲則¹

Yoshihiko Hayashi¹ Shinya Fujie^{2,1} Ishin Fukuoka¹ Hiroaki Takatsu¹ Tetsunori Kobayashi¹

¹ 早稲田大学 ² 千葉工業大学

¹ Waseda University ² Chiba Institute of Technology

Abstract: Passive information consumption would an adequate type of information behavior for receiving the content of, for example, a news article. It may however be boring in many cases and even painful in some cases, especially when the information content is delivered by employing speech media. The user of a speech-based information delivery system, for example a text-to-speech system, usually cannot interrupt the ongoing information flow, inhibiting her/him to confirm some part of the content, or to pose an inquiry for further information seeking. We thus argue that spoken dialogue is a suitable media for enabling interactive information access that coordinate passive information consumption and active information seeking. This paper shows that a carefully designed spoken dialog system could remedy these undesirable situations, and further enables an enjoyable conversation with the users. The key technologies to realize such an attractive speech-based interactive information access system are: (1) pre-compilation of a dialog plan based on the analysis of a source content, and (2) the dynamic recognition of user's state of understanding and interests during the course of conversation. This paper illustrates technical views to implement these functionalities, and discusses a dialog example to exemplify our approach.

1 はじめに

人間の情報に関する行動 (information behavior) のうち、情報獲得・収集 (information acquisition) に関する行動は大きく、意図的な情報探索 (intentional information seeking) と、意図性のない受動的な情報行動 (unintentional passive information behaviors) に分けられるとされ [1]、情報学の分野では、主に前者を導く動機や状況に関するモデルの研究が行われてきた [3]。

コンピュータサイエンスの領域においても、その焦点はもちろん前者にあり、情報検索 (information retrieval)、あるいは少し広い概念としての情報アクセス (information access) のシステムについて、様々な観点からの研究開発が活発に行われてきた。

以上の研究状況の背景を推察するに、情報遭遇 (information encountering) などの受動的な情報行動は、主として偶発的な状況によることから、研究的な要素に乏しいと考えられてきたのではないかと考えられる。

しかしながら、我々の日常の情報行動の実際をみれば、両者の区別は必ずしも明白ではなく、むしろ、これらの情報行動の状態を自由に遷移する過程であると考えるのが妥当であろう [2]。

さて本研究では、音声対話によるニュース記事アクセスシステムをとりあげる。ユーザ側からみれば、ニュースに関する情報を音声メディアを用いて獲得し、あるいは、消費する情報アクセスシステムであるが、システム側の観点から言えば、音声メディアを用いて、ユーザに伝えたい・伝えるべきニュースを伝達するという情報伝達システムである。

システムが一方向的に記事の内容を読み上げるとすれば、ユーザは読み上げ音声を黙って聞き続ける必要がある。ユーザにとって内容的に冗長である可能性もあるし、そもそも記事の内容に興味がないことに気づく場合もあるだろう。

このような情報提示システムの対極に、記事に関する簡単な内容 (例えば記事の見出し) を与えた後に、ユーザからの質問を一問一答形式で受け付けるモードに移行する質問応答型のシステムが考えられる。このよう

*連絡先: 早稲田大学理工学術院 実体情報学博士プログラム
〒169-0072 新宿区大久保 2-4-12 ラムダックスビル 3F
E-mail: yshk.hayashi@aoni.waseda.jp

S_1: 羽生結弦選手が
 U_1: うん
 S_2: 国際大会を欠場することになったよ
 U_2: え?
 S_3: 欠場するんだ、腰の痛みのためだって
 U_3: 腰の痛み... って?
 S_4: 練習中に腰を痛めたということなんだ
 ...

図 1: 想定する対話の断片例. “S_n:” はシステム発話, “U_n:” はユーザ発話を表す.

なシステムのユーザは, 適度な量の情報を得るまで, 質問を発し続けることが必要になる.

先の情報獲得における意図性の議論からすれば, ユーザは両者のモードを, その「状況」に応じて, しかも簡単な手段によって, 行き来できることが望まれる. そこで, 本研究が想定するような対話を単純化した断片の例を図 1 に示す. システムから伝達された情報に対してユーザは, 必ずしも言語的ではない即応的な反応 (U_1=肯定的, U_2=疑問) によって理解状況を示したり, さらに対話の過程で生じた情報要求 (国際大会を欠場する理由) をある程度明確な言語表現を用いて示し (U_3=問い返し) たりする. システムは, 必要に応じてこれらのユーザの状況を推定し, 適切と思われる情報を付加しながら応答を返す (S_4).

本研究の前提, あるいは, 主張は, このような受動的な情報獲得を主体としつつもインタラクティブ性を要する・有する情報行動の支援形態として, 音声対話が適しているという点にある.

本稿の以下では, まず情報学における関連研究を参照しながら, 上記の議論を補強し, 本研究のスタンスを明確化する. 次に, 現在開発中のニュース記事を対象とする音声対話システムについて述べ, 最後に今後の課題や研究の方向性について論じる. なお, 音声対話システムに関する内容は, 当研究グループにおける既発表 [17] の内容によっている.

2 情報行動における受動性と能動性

これまでに提案されている情報行動の分類には様々なものがあるが, Erdelez による分類 [2] を図 2 に示す. ここでは, 非意図的な情報行動は機会主義的情報獲得 (Opportunistic Acquisition of Information: OAI) と呼ばれており, その主な下位分類として, 情報遭遇 (information encountering) が位置づけられている.

Erdelez はさらに, (1) 気づき (noticing); (2) 停止 (stopping); (3) 検討 (examining); (4) 獲得 (captur-

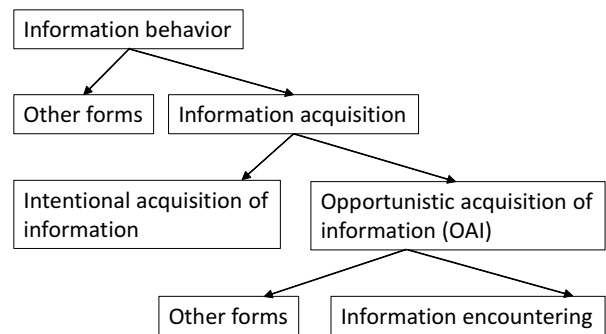


図 2: Erdelez による情報行動の分類 ([2] より作図).

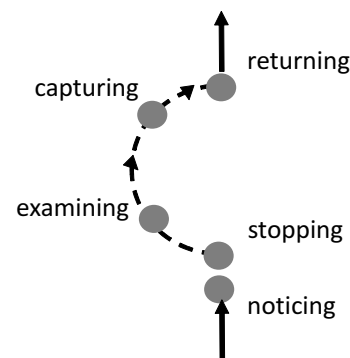


図 3: Erdelez による情報遭遇の機能モデル. 実線部が foreground interest, 点線部が background interest を表す. ([2] より作図).

ing); (5) 復帰 (returning). の各段階からなる情報遭遇における機能モデル (図 3) を提示した.

このモデルでは, ユーザは彼/彼女の主要な関心 (foreground interest) に関わる能動的な情報探索タスクを実行していることが仮定されているが, 重要なことは, ユーザはこの過程の中で関連する関心 (background interest) に気づき, foreground の情報探索タスクを一旦停止したうえで, background に関する情報行動 (検討・獲得) を行い, その後に foreground の情報探索タスクに復帰するという点である.

以下, 図 1 に示した対話例をこのモデルをと照らし合わせて考えてみる. この対話における「主要な関心」は, 対象のニュース記事により定まる「羽生結弦選手の国際大会欠場」にある. これが主要な関心となる契機については問わない¹が, システムが記事内容の伝達を行っている間は, ユーザは基本的には受動的な情報消費のモードにある.

本研究では, この対話でのユーザによる U_3 の発

¹すなわち, システム側が勝手に見つけた記事 (受動的なあるいは機会主義的な情報獲得) かもしれないし, ユーザによるある種の情報検索の結果として選択されたもの (意図的な情報探索) であってもよい.

話(「腰の痛み... って?」)の背景には、図3のモデルにおける noticing に相当する過程があると考えられる。すなわち、foreground に対する情報行動の過程の中で background に対する情報行動へのシフトが起こっている。ただし、図3の Erdelez のモデルでは、能動的で意図的な情報探索の過程において偶発的な情報遭遇が生じているのに対し、図1に示す対話においては、ある程度の明示性のある言語表現(「腰の痛み... って?」)によって、モードのシフトが起こっている点が異なる。しかしながら、background に関する情報行動のモードが一段落した後は、foreground の情報行動のモードに復帰することは共通している。

以上にみたような、ユーザにとって自然で効率の良い情報行動は、受動的な情報消費から、意図的な情報要求による情報探索までの情報行動のモードをその「状況」に応じて、しかも簡単な手段によって、行き来することにより達成される。本研究では、このような情報アクセス・情報伝達を実現するために音声対話が自然で効率よいメディアであるという前提にたち、次節で述べるような音声対話によるニュース記事アクセスシステムを提案する。

3 音声対話によるニュース記事アクセスシステム

3.1 システムの要件

理想的な対話システムの実現へ向けては様々な課題が存在するが、少なくとも以下のような要件を考慮する必要がある。

伝えるべきニュース記事の選択: どのようなニューストピックを対象とするかの決定は、本研究の範囲外とする。すなわち、「本日の重大ニュース」でも、「本日のおすすめニュース」でも、ユーザによる情報検索の結果として選択されたものでも良い。

伝えるべき内容の選択と構成: システムは、対象とするニュース記事(群)が与えられたとき、最低限どのような内容を伝えるべきかを決定する必要がある。ここでいう最低限伝えるべき内容とは、対話においてユーザが完全に受動的である場合においても、システム側がとにかく伝達しようとする骨格的な情報内容である。要約と言っても良いだろう。このような要約に相当する情報内容を補足する補助的な内容は、ユーザからの具体的な情報要求に基づいて提示することになる。

ユーザの状況の把握と対応: ユーザは、システムの話に対する自身の理解状況や伝達内容への興味の状況

を反映して、肯定的・否定的な短い即応的な発話(以下、即応的情報反応と呼ぶ)や、もう少し明示的な情報要求を発することが想定・期待されている。したがって、システムはこれらのユーザの反応・発話からユーザの状況を適切に把握し、さらにはそれに応じた応答を返す必要がある。

リズムのある対話の実現: リズムのある対話を実現するためには、即応的情報反応を含むユーザの短い発話に対して、システムは素早く応答できることが望まれる。よって、対話の過程で提示されるユーザからの情報要求をある程度見越して、「こう聞かれたら、こう答える」ということを定めておくことが必要となる。

3.2 システムの構成

以上のような要件を(ある程度)満たすものとして、我々が提案する音声対話システムの構成を図4に示す。システムは大きく分けて、ニュース記事をもとに対話に利用する発話計画を生成する事前処理部と、ユーザを相手に対話を行う対話システム部の二つから構成される。

事前処理部は、インターネットからの取得などによって与えられるニュース記事を解析して構造化する構造解析部と、その結果をもとにユーザの反応を織り込んだシステム発話計画を生成する計画生成部からなる。発話計画は、記事における主要な内容を伝達するための主計画と、それを補足する補助的な内容を伝達するための副計画からなる²。構造解析部については3.4で、計画生成部については3.5でそれぞれ詳細を述べる。生成された発話計画は、発話計画データベースに保存される。

対話システム部は、発話計画を読み込み、それに従って対話を進める。音声認識器は、ユーザの短い反応を認識する。対話制御部は、発話計画に従ってシステムの発話内容を含む発話文を音声合成器に出力する。また、システムの発話に対するユーザの反応に応じて、事前に生成した発話計画に従って発話内容の調整を行う。音声合成器は、対話制御部から生成された発話文を音声に変換してユーザに提示する。

本システムにおいて、音声認識器はATR-Trek製のものを使用している。また、音声合成器はOpen JTalk³を基に開発したものを採用している。

3.3 発話計画とユーザの反応・応答

ここでは、システム側の視点から発話計画とユーザの応答の関係について述べる。本システムでは、想定

²これらはさらにネストしていてもよい。

³<http://open-jtalk.sp.nitech.ac.jp/>

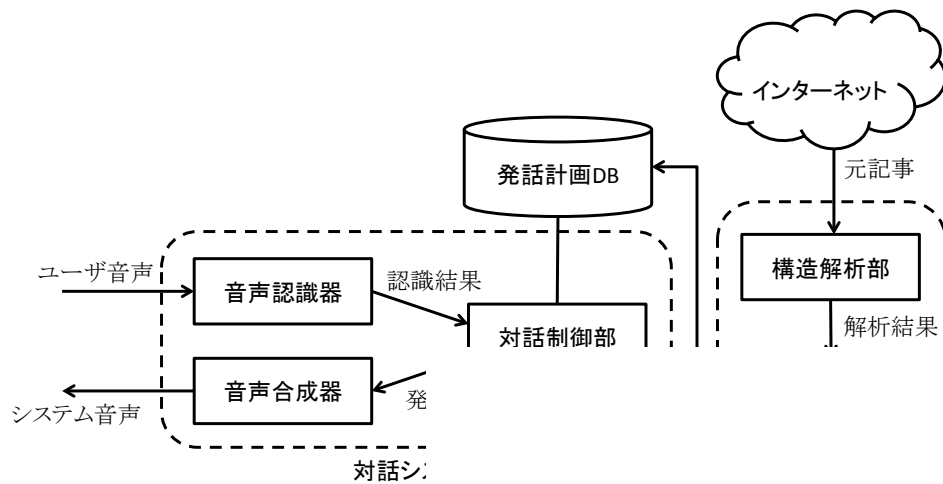


図 4:

されるユーザの反応を織り込んだ形で発話計画を事に生成しておき、実際の対話時にはそれに従うこと効率的な情報伝達を実現する。ここではその発話計画がどのようなデータ構造を持つか、また、それに従って対話制御部がどのように対話を進めていくかについて述べる。

まず、発話計画のデータ構造を図5を用いて説明する。図に示した通り、発話計画は状態遷移構造として表される。各アークに示された $U_{i,j}$ はユーザ発話をし、 $S_{i,j}$ は、システム発話を表す。ユーザ発話は音声認識部へ送られ、システム発話はそのまます成部へ渡される。

ユーザ発話 $U_{i,j}$ は即応的情報反応を含む短い応答を想定しており、下記の2つのカテゴリのうちいずれかに分類される。

肯定的応答 (ACK) 「うん」「へー」といった相槌など、システムの発話進行に肯定的な態度を表す反応。システムの発話の一部を下がり口調のイントネーションにより反復する場合を含む

否定的応答 (NACK) 「え?」といった、システムの発話進行に否定的な態度を表す反応。上がり口調のイントネーションによるシステム発話の一部の反復を含む

すでに述べたように、システム発話の一部を反復することは、ある程度明確な情報要求がユーザにおいて発現している状況、すなわち、情報行動のモードにシフトが起こっている状況を示唆する。また、否定的な応答は、システム発話のいずれかの部分が聞き取れなかった状況、あるいは、発話された内容に対して明示的でない情報要求が生じている状況を表していると考えられる。

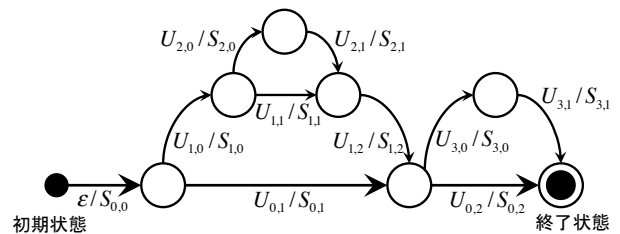


図 5: 発話計画の構造。 $U_{i,j}$ はユーザの反応、 $S_{i,j}$ はシステム発話を表す。

ϵ はユーザの反応を見ずに状態が遷移することを表す。図中で太いアークにより表されている部分を主計画と呼ぶ。主計画は、システムが最低限伝えるべき、記事の骨格をなす内容情報を表す。ユーザが受動的な態度(すなわち、反応がないか、ACKのみを示す)を取り続ける限り、システムは主計画に従って淡々と、この内容情報の伝達を進めることになる。

本システムにおける対話制御は非常にシンプルで、アークに与えられたシステム発話 $S_{i,j}$ の内容を音声合成器に出力した上で状態を遷移させる。各状態ではユーザの反応を待ち、その内容に従って次の遷移を行う。例えば主計画上にいる際は、特にユーザの反応が得られなくてもそのまま発話を続けることが好ましいと思われるので、図中、 $U_{0,1}$ や $U_{0,2}$ などは、特に反応が無くても時間経過(例えば0.6秒無反応で経過)によって遷移させる。

否定的応答が認識された場合(例えば $U_{3,0}$ が得られた場合)は、その直前のシステム発話 ($S_{0,1}$, あるいは $S_{1,2}$) を補足する副計画に従い、情報を提示する発話 ($S_{3,0}$) を生成する。

このように、想定されるユーザの反応を織り込んだ計画によって対話制御を行うことで、素早く効率的に

ソチオリンピック、／フィギュアスケート男子の／金メダリスト、／羽生結弦選手が／腰の／痛みの／ため、／今シーズンの／初戦と／して、／来月フィンランドで／出場を／予定していた／国際大会を／欠場することになりました。

図 6: 文節単位に分割された文

ユーザの理解や知識に見合った情報を伝達することが可能となる。

なお、ユーザの反応として、上記のカテゴリに含まれない、より明示的な質問により情報要求が提示された場合は、例外処理として質問応答型の対話制御に一時的に切り替えることを考えているが、その詳細は現在検討中である。

3.4 構造解析部

構造解析部では、対象とするニュース記事をもとに、発話計画を生成するために必要な情報を構造化する。すでに複数記事を対象とする場合の検討を進めているが、本稿では単一の記事を対象とする。

構造解析の目的は、ニュース記事が持つ情報をもとに発話計画を立てるための情報を抽出することにある。すでに示したように、発話計画はニュースの要点を伝える主計画とそれを補う副計画からなるので、構造解析部では主計画に含めるべき情報(以降、主情報)と、周辺情報と主情報の関係性を抽出するという課題がある。

構造化: 構造解析は、文節単位の係り受け関係をもとに行う。そのため、まず文を文節単位に分割し、係り受け解析を行う。本研究では、形態素解析と係り受け解析に、Juman⁴、KNP⁵をそれぞれ利用した。

例として、ウェブニュース記事⁶の一文を文節に分割した例を図6に示す。さらに、係り受け解析に基づいて、文節をノードとする依存構造木を作成する。図7に、図6をもとに生成した依存構造木を示す。図中では省略しているが、各アークは係り受けの関係属性を保持し、各ノードは当該の文節の文法的情報を持つ。例えば、「羽生結弦選手が」に対応するノードは、「人物、主題」という情報を持つ。

主情報の抽出: 図7中、太枠で囲まれた文節がこの文における主情報であり、主発話計画を構成する。主情報として、まず、対象文の主辞となる文末の述語文節と、それに対する必須格要素となる文節を選択する。次に

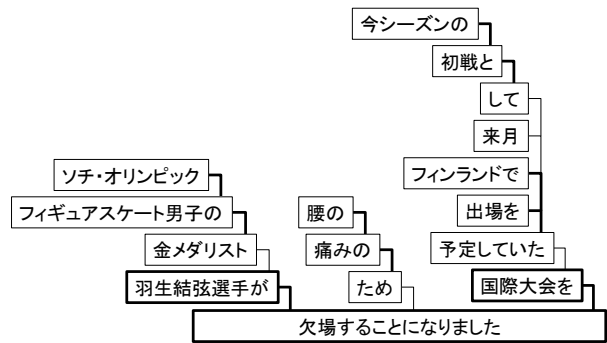


図 7: 図 6 の文に対する依存構造木

これらの文節に対して、次に説明する省略不可避性を有する文節を抽出する。図中では、太線によるエッジが省略不可避性を表す。

省略不可避性とは、「当該文節の係り先の文節が発話される場合は、当該文節が省略されてはならないこと」を表す。例えば「ため」という形式名詞からなる文節は、それだけでは特定の意味を持たないため、「痛みの」を省略できない。また、「痛みの」に対しての「腰の」が省略不可避であるかは微妙であるが、本研究では、野本による統計的な文圧縮の研究 [15]における依存構造木の「刈り込み」に準じた方法により、省略不可避性の判定を行う。主情報は主計画に組み込まれるため、主情報として選ばれた文節から文の最後の述語となる文節にたどり着くために通る経路上の文節の係り受け関係には、全て省略不可避性があるとする。

以上は、言わば非文法的な文の発話を回避するために必要な処理であるが、さらに、内容的に含めたほうが良いと思われる重要語も主情報に含める。このために、松尾らの手法 [18] をニュース記事の性質を踏まえて変更した重要語抽出処理を用いている [17]。

3.5 計画生成部

計画生成部では、構造解析部で得られた結果を用いて、3.3で述べた発話計画を作成する。ここでは、3.4での例をもとに生成した、図8に示した発話計画の例をもとに、主計画、副計画の生成について説明する。

主計画の生成: 主計画は、前節で述べた主情報により構成する。例に挙げた文では、「羽生結弦選手が」、「国際大会を」、「欠場することになりました」という3つの文節が主情報であった。

選ばれた文節をもとに語順にしたがって配列して作成した文を適切な長さに分割する。これは、文節を連結した文をそのまま読み上げると、ユーザが短い反応を挟む間を奪ってしまう可能性があるためであ

⁴<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

⁵<http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

⁶NHK NEWSWEB, <http://www3.nhk.or.jp/news/>

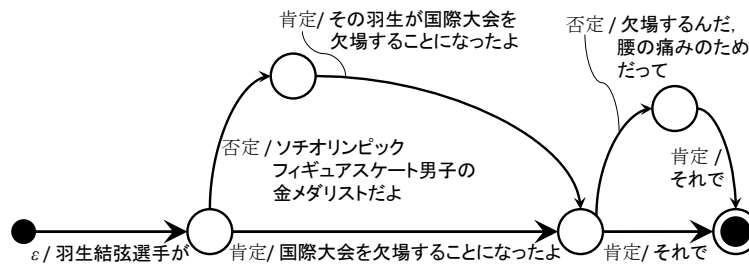


図 8: 発話計画の例

る。一方で、あまり短い単位毎に発話を区切ってユーザの反応をうかがうのは、ユーザに煩わしさを感じさせる可能性があり適切ではない。このため、簡単な文法的な規則を適用し、文節を連結することで適切な長さの発話を構成する。ここでは、海木らのポーズ挿入規則 [9] を参考に、次の条件を満たす箇所では分割し、それ以外では連結する。

- 当該文節の句が右枝分かれ、かつ、先行文節が左枝分かれ
- 当該文節が読点 (、) を含む

この規則を先の例に適用すると、「羽生結弦が」と「国際大会を欠場することになりました」という2つの発話となる。つまり、それぞれの区間をシステムが発話した後は一定のポーズがおかれるので、ユーザが反応を入れやすくなると考える。

副計画の生成 構造化された情報の中で、主計画に含まれる文節に係っている文節のうち、省略不可避性がない文節が副計画を構成する。さらに主計画の場合と同様に、各文節から省略不可避性を有する文節を再帰的に辿って取得した文節を計画に含める。図7の例で、「国際大会を」に対する副計画を生成するとき、まず「国際大会を」に係る文節である「予定していた」が選択され、それに対する省略不可避性をもとに「出場を」「フィンランドで」が選択される。全体として「フィンランドで」「出場を」「予定していた」という文節が「国際大会を」の副計画を構成することになる。ここで選択されなかった「来月」などの文節は、「予定していた」に対する副計画となる。このように発話計画は階層性を持つ。副計画も、主計画と同様の規則で文の分割を行う。

4 対話例と課題

提案システムとユーザの対話例を図9に示す。システムが主計画に沿って発話をするので、ユーザは相槌をはさむ(肯定的応答 ACK)、あるいは黙っている

S: 羽生結弦選手が	[主計画]
U: うん	[肯定応答]
S: 国際大会を欠場することになったよ	[主計画]
U: え?	[否定応答]
S: 欠場するんだ、腰の痛みのためだって	[副計画]
U: へー	[肯定応答]
S: それで、日本スケート連盟によると	[主計画]
U: うん	[肯定応答]
S: 羽生選手は全治4週間と診断されたんだって	[主計画]
U: そうなんだ	[肯定応答]
S: スケート連盟は	[主計画]
S: 『症状は重くない。練習が全くできない状況ではない。中国大会に向けて	[主計画]
U: 中国大会?	[反復応答]
S: うん、グランプリシリーズのだよ	[副計画]
U: ふうん	[肯定応答]
S: それで、万全の状態を望むためだ』としているよ	[主計画]
U: そうなんだ	[肯定応答]

図 9: 対話例。“S:”はシステム発話，“U:”はユーザ発話を表す。システム発話には主計画と副計画のどちらから生成されたか、ユーザ発話には反応がどの応答に分類されたか、をそれぞれ付与している。

だけでニュースの要点が得られる(受動的な情報消費)。また、ユーザの知らない言葉、興味を引いた単語、あるいは、理解できない表現が出て来たときには相槌(否定的応答 NACK)や聞き返しを行うことが想定される。このときシステムは、対応した副計画に沿った発話を生成することで情報を補足する(backgroundの情報行動)。このように、ユーザの状況(理解や興味の状態)に合わせてニュース記事の内容を音声によって効率よく伝達することが実現できる。

むしろ、課題も多く残されている。例えば、「国際大会を」と「欠場することになった」は、計画上では一つの発話としてまとめられている。二つの情報を一つ

にまとめた発話に対してユーザの聞き返しがあった場合、どちらに関する補足を行うかは自明ではない。また、一つの情報に複数の補足情報が存在することもある。従って、補足対象が定まったとしても、どの補足情報を伝えるべきかは別途決めなければならない。現状ではこれらの問題に対して、「発話中で最も後ろの情報に対して優先的に補足を行う」「時間や場所以外の補足情報を優先する」といった規則を適用しているが、より適切な情報提示のためには、どの補足情報が重要かといったことを考慮すべきである。また、用語や人物の説明といった補足情報は、それらの一般的な知名度や、個人の知識や嗜好などによって変化することを踏まえ、ユーザの反応に対する適切な補足情報の提示となるような発話計画を立てる手法を確立することが求められる。

5 関連研究

従来より、特定のタスク(交通案内や天気情報提供など)を対象として、ユーザからの発話に応じて情報を提示する質問応答型の対話システムが研究されてきた[8, 10, 14]。このようなシステムでは、ユーザの明示的な情報要求に応じて限定的な情報を確実に提供することに主眼が置かれていた。

近年、質問応答と組み合わせ、システム側からユーザに主体的に情報を提示する対話システムの研究も進み[19, 11, 12]、文書で表わされるような、まとまった量の情報提供を行う音声対話システムも提案されてきた[7, 20]。しかしながら、これらのシステムも基本的にはユーザの質問に対してシステムが回答を提示するという点では質問応答型の対話になっていると言いたい。

ところで、ユーザにとって質問という行為は、システムの発話内容を理解した上で、問いかける内容を明示的に言語化する必要があるため、比較的負荷が高い。そのため、この種のシステムにおいてはユーザからの質問がなされにくく、システムは要約が提示することが主な機能となり、対話をとおして必要十分な情報を効率良く伝達することは困難であった。

これを解決するには、ユーザが質問を発しやすい状況を作り出すことが必要になる。音声対話の特性を考えれば、ここで言う質問には、言語表現を用いた明示的な情報要求だけでなく、相槌や聞き返し、相手の発話の一部の反復などの短い反応(本稿では即応の情報反応と呼んだ)を含めて考えるべきであることはすでに論じたとおりである。これらの反応や情報要求を認識するには、発話される語句の音声認識が必要であることは当然であるが、状況や態度を表出する手段としてのパラ言語情報の識別が重要となる。さらには、身

振り手振りや顔の動きなどのマルチモーダル情報も手がかりを与える。当研究グループでは、これまでもユーザの短い反応を韻律情報などを利用して認識するシステムを提案してきている[16, 4, 6]。

6 議論

適用領域: まず、提案したような音声対話による情報アクセス・情報伝達システムの適用領域についてであるが、やはり、音声対話メディアの特性から、いわゆるユーザが「手が離せない」状況が考えられる。このような条件に合致した適用場面としては、機器の運転中や、料理などの手作業中などが考えられる。一方、システムに対して音声で話しかけたり反応を返したりが行いやすいかという点も問題になる。この問題には、もちろんユーザの嗜好や特性も影響するが、ある種の擬人化エージェント的なインタフェースが有用である可能性も考えられる。また、ユーザの反応を引き出しやすいような、システム発話の生成[5]も有効な要素となろう。

情報アクセスシステムとしての位置づけ: 通常のテキストを中心とする視覚的メディアを用いた情報アクセスの研究は盛んに行われている。しかしながら、いわゆるサーチエンジンを超えるようなポピュラリティを得ているシステム・インタフェースはほとんどないと言える。一方で、先に指摘したように、音声メディアが適している、あるいは、音声メディアしか使えないような利用状況が考えられる。ただし、記事や文書のようなまとまった情報を音声で伝達したり、ブラウジングすることには困難がある。そのような意味で、音声対話によってもたらされるインタラクションを導入することにより、「基本は受動的だけど、能動的なつまみもできる」情報アクセスシステムを実現しようとする本研究の方向性には、これまでにはなかった可能性があると考えられる。

対話システムとしての位置づけ: 対話システムの分類の軸として、対話の主導権をシステムが持つか、ユーザが持つか、または両者の混合かというものがある[13]。本研究の音声対話システムは、現在の範囲においては、「基本はシステム主導で、必要に応じてユーザ主導」になる。ただしこれは、微妙なコントロールを短い音声反応で行える範囲に限定しての話である。この成約は、対話制御部を非常にシンプルなものにすることに貢献するが、今後もっと明示的で複雑なユーザの情報要求を扱おうとする場合、情報要求をシステム内部の情報検索過程に対応付けるための対話や、答えられない要求に対する対話などの複雑な対話制御が必要となる。もっ

とも、サービス・機能的な側面から、どの程度のことまで行おうべきか、行えるかを定めるための検討も必要となる。

7 おわりに

ニュース記事から、ユーザの反応を想定した発話計画を作成し、それに従って対話を行うことで、ユーザから見れば必要十分な情報アクセス、システムから見れば効率的な情報伝達が行える音声対話システムを提案した。提案システムのユーザは、音声対話において自然と考えられる反応を返したり、情報要求を提示することにより、過不足のない情報アクセスを実現することができる。

今後は提案システムの枠組みを発展させ、さらに効率的で豊かなインタラクションを実現することを目指す。そのための課題として、パラ言語を利用したユーザの反応の認識、複数のニュース記事群を対象とした、より適切な情報内容の選択と構成、ユーザの反応を引き出すような対話の展開、親しみやすい発話音声の生成など個々の技術における精度や使い勝手の向上が挙げられる。

本システムは明確なタスク達成指向のシステムではなく、また、話を継続することを目的とする雑談システムでもない。また、今後は楽しく役立つ情報対話といった要素も加味していきたいと考えている。その意味で、評価の観点や方法論が現状では未確立である。よって、構築したシステムを実際の対話で評価しながら、これらを確認することも課題である。

参考文献

- [1] D.O. Case: *Looking for Information, A Survey of Research on Information Seeking, Needs, and Behavior, Second Edition*, Academic Press (2007)
- [2] S. Erdelez: Information encountering, In [3], pp.179–184, (2005)
- [3] K.E. Fisher, S. Erdelez, and L.E.F. EmKechinie (Eds), *Theories of Information Behavior*, Information Today, Inc. (2005)
- [4] S. Fujie, R. Miyake, and T. Kobayashi: Spoken dialogue system using recognition of user's feedback for rhythmic dialogue, *Proc. Int. Conf. Speech Prosody*, OS2-4 (2006)
- [5] K. Iwata and T. Kobayashi: Speaker's intentions conveyed to listeners by sentence-final particles and their intonations in Japanese conversational speech, *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 6895–6899 (2013)
- [6] T. Kobayashi and S. Fujie: Conversational Robots: An Approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting, *Acoust.Sci. & Tech.*, Vol. 34, No. 2, pp. 64–72 (2013)
- [7] Y. C. Pan, H. Y. Lee, L. S. Lee: Interactive spoken document retrieval with suggested key terms ranked by a Markov decision process, *IEEE Trans. Audio, Speech, and Language Processing*, Vol. 20, No. 2, pp. 632–645 (2012)
- [8] S. Seneff, J. Polifroni: Dialogue management in the Mercury flight reservation system, *Proc. 2000 ANLP/NAACL Workshop on Conversational systems*, Vol. 3, pp. 11–16 (2000)
- [9] 海木延佳, 匂坂芳典: 局所的な句構造によるポーズ挿入規則化の検討, *信学論 (D-II)*, Vol. J79-D-II, No. 9, pp. 1455–1463 (1996)
- [10] 駒谷和範, 上野晋一, 河原達也, 奥乃 博: ユーザモデルを導入したバス運行情報案内システムの実験的評価, *情処学研報, SLP*, 2003.75, pp. 59–64 (2003)
- [11] 杉山 聡, 堂坂浩二, 川端 豪: 音声対話によるテキスト内容の伝達方法, *情処学論*, Vol. 41, No. 6, pp. 1883–1894 (2000)
- [12] 杉山弘晃, 南 泰浩: 情報提示対話を主導するシステムのためのユーザの潜在的情報要求の推定, *信学論 (A)*, Vol. J95-A, No. 1, pp. 74–84 (2012)
- [13] 中野幹生, 駒谷和範, 船越孝太郎, 中野有紀子: 対話システム, コロナ社 (2015)
- [14] 西村良太, 北岡教英, 中川聖一: 応答タイミングを考慮した雑談音声対話システム, *人工知能学研資, 言語・音声理解と対話処理研究会*, Vol. 46, pp. 21–26 (2006)
- [15] 野本忠司: 係り受け構造の刈り込みと CRF による文の要約, *言語処理学会年次大会*, pp. 488–491 (2008)
- [16] 藤江真也, 江尻 康, 菊池英明, 小林哲則: 肯定的/否定的発話態度の認識とその音声対話システムへの応用, *信学論 (D-II)*, Vol. J88-D-II, No. 3, pp. 489–498 (2005)
- [17] 藤江真也, 福岡維新, 麥田愛純, 高津弘明, 林 良彦, 小林哲則: 効率的な情報伝達を志向した音声対話システムの提案, *人工知能学会 第 74 回 言語・音声理解と対話処理研究会, SIG-SLUD-B501-02*. (2015)
- [18] 松尾 豊, 石塚 満: 語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム, *人工知能学論*, Vol. 17, No. 3, pp. 217–223 (2002)
- [19] 翠 輝久, 河原達也, 正司哲朗, 美濃導彦: 質問応答・情報推薦機能を備えた音声による情報案内システム, *情処学論*, Vol. 48, No. 12, pp. 3602–3611 (2007)
- [20] 吉野幸一郎, 河原達也: ユーザの焦点に適応的な雑談型音声情報案内システム, *人工知能学研資, 言語・音声理解と対話処理研究会*, Vol. 70, pp. 53–58 (2014)

アクティビティ評価情報とNMFの融合による ソーシャルメディアの信頼リンク予測

Combining Activity-evaluation Information with NMF for Trust-link Prediction in Social Media

松谷 貫司¹ 熊野 雅仁¹ 木村 昌弘^{1*}
Kanji Matsutani¹ Masahito Kumano¹ Masahiro Kimura¹
齊藤 和巳² 大原 剛三³ 元田 浩⁴
Kazumi Saito² Kouzou Ohara³ Hiroshi Motoda⁴

¹ 龍谷大学理工学部電子情報学科

¹ Department of Electronics and Informatics, Ryukoku University

² 静岡県立大学経営情報学部

² School of Administration and Informatics, University of Shizuoka

³ 青山学院大学理工学部情報テクノロジー学科

³ Department of Integrated Information Technology, Aoyama Gakuin University

⁴ 大阪大学産業科学研究所

⁴ Institute of Scientific and Industrial Research, Osaka University

Abstract: Acquiring networks of trust relations among users in social media sites such as item-review sites is important for analyzing users' behaviour and efficiently finding reliable information on the Web. For an item-review site, we address the problem of predicting trust-links among users. Recently, non-negative matrix factorization (NMF) methods have been shown to be useful for trust-link prediction in such an site, where both the link and activity information is employed. Here, a user activity in an item-review site means posting a review and giving a rating for an item. Aiming to improve NMF methods for trust-link prediction, in this paper, we propose such an NMF method that incorporates information of people's evaluations for users' activities as well as information of trust-links and users' activities. Also, we apply it to an analysis of users' behaviour. Using real data of an item-review site, we experimentally demonstrate the effectiveness of the proposed method.

1 はじめに

近年, Digg, eBay, Epinions, Facebook などのソーシャルメディアの普及に伴い, オンライン世界に大規模ネットワークが生成されている. ソーシャルメディアユーザの間の信頼ネットワークはソーシャルネットワークの一種であり, Web 上で信頼できる情報を効良く発見することに役立つ. ソーシャルメディアサイトに投稿されたニュースや意見は, ソーシャルネットワークを介して急速かつ広域的に拡散し, 多数の人々に共有されるようになる. このように, 信頼ネットワークは人々の日々の生活の中で重要な役割を果たす. したがって, 信頼ネッ

トワーク (ソーシャルネットワーク) のマイニングに対して, 情報拡散分析 [1], [2], [3], [4], [5], [6], [7] や, ソーシャルリンク予測 [8], [9], [10], [11], [12], [13], [14] などの様々な研究が行われている.

ソーシャルメディアサイト上のユーザの行動を分析するためには, そのサイトにおけるユーザ間の信頼関係により構築されるネットワークを獲得することが必要不可欠である. しかし, 信頼ネットワークは時間経過と共に継続的に進化し, かつ, プライバシーの問題も存在することから, 分析に必要な特定の時点についての完全な信頼ネットワーク構造を獲得することは一般に困難である. また, オンライン世界において, 人々の行動パターンや嗜好は長期間では大きく変化すると考えられる. そこで, 本論文では, 直近の過去において活動的なユーザを対象に, 直近の未来に生成される信頼リンクを予測す

*連絡先: 〒520-2194 大津市瀬田大江町横谷 1-5
龍谷大学理工学部電子情報学科
E-mail: kimura@rins.ryukoku.ac.jp

る問題を論じ、この問題の効率的な解法を提案する。ここに、我々の提案法は未観測リンクの予測問題にも適用可能であることに注意しておく。

多くのソーシャルメディアサイトでは、信頼リンク生成機能とともに、アクティビティ実行機能が提供されており、ユーザは与えられたアクティビティ群の中から一つを選択して実行することができる。たとえば、アイテムレビューサイトにおいて、ユーザは与えられたアイテム群の各アイテムに対してレビューの投稿および評点付けを行うことができる。信頼リンクとアクティビティを提供するソーシャルメディアサイトにおける信頼リンク予測問題に対して、Tang ら [14] はリンクとアクティビティの情報を利用し、類同性 (*homophily*) と呼ばれる社会学の理論を利用した非負値行列分解 (NMF) 法を提案した。ここに、類同性理論では類似するユーザ (類似するアクティビティを実行したユーザ) ほど信頼関係を確立する可能性が高いことが示唆されている [10], [15]。 *hTrust* と呼ばれる彼らの手法は NMF に基づいており、NMF 法 [16] は協調フィルタリング、文書クラスタリング、リンク予測などの多くの応用に対して有効であることが示されている。彼らは、*hTrust* が従来法より高性能であることをアイテムレビューサイトの実データを用いた実験で示した [14]。

Tang ら [14] と同様に、我々もアイテムレビューサイトに焦点を当てる。最近の多くのアイテムレビューサイトでは、あるユーザのアイテムへのレビューに対して、別のユーザがそれを有益であると思ったときに感謝メッセージを投稿することができる。信頼リンク予測に対する NMF の性能向上のために、我々はそのような情報と NMF との融合を考える。本論文では、信頼リンク予測問題に対して、信頼リンクとユーザアクティビティの情報とともに、ユーザアクティビティの評価情報を組み込んだ新たな NMF を提案し、それをソーシャルメディアサイトにおけるユーザ行動の分析に応用する。また、アイテムレビューサイトの実データを用いて提案法を評価する。まず、データセットの統計解析を行い、特に、受け取った感謝メッセージ数と信頼リンクの入次数との間に相関があることを示す。次に、信頼リンク予測問題に対して、提案法が *hTrust* とそれに類似した手法に比べて高性能であることを実証する。さらに、信頼リンク生成の観点から、データセットにおけるユーザ行動を分析する。

本論文の構成は以下の通りである。2 節では、信頼リンク予測問題を定式化し、信頼リンク予測に対してリンク情報とアクティビティ情報を用いた NMF 法を説明する。3 節では、提案 NMF 法を示し、4 節において評価実験の結果を報告する。5 節では、主たる結果を要約し、本論文の結論を述べる。

2 NMF による信頼リンク予測

Tang ら [14] はソーシャルメディアにおける信頼リンク予測に対して、リンク情報とアクティビティ情報を用いた非負値行列分解 (NMF) 法である *hTrust* を提案し、その有効性を示した。また、類似する NMF 法が Zhu ら [17] によって提案されている。彼らは Web ページの分類に対してリンクと文書の情報を用いた NMF 法である *Joint Link-Content Matrix Factorization (JLCMF)* を提案した。本節では、ソーシャルメディアにおける我々の信頼リンク問題を定式化したのち、*hTrust* と *JLCMF* を再訪する。

2.1 問題の定式化

本論文では、信頼リンクとアクティビティを提供するソーシャルメディアサイトに対して、直近の過去において活動的なユーザを対象に、直近の未来に生成される信頼リンクを予測する問題を扱う。ここに、このようなサイトにおけるユーザアクティビティとは、与えられたアイテム群の各アイテムに対してレビュー投稿および評点付けを行うこととする。さらに、サイト内のアイテムにユーザが行うレビューに対して、別のユーザがそれを有益であると思ったときに感謝メッセージを投稿することができる。つまり、ユーザアクティビティの評価情報をサイト内にて観測することができる。

我々は、直近の Δt_0 か月内 (観測期間 I_0 と呼ぶ) に少なくとも n_1 本の信頼リンクを結び、かつ、少なくとも n_2 本の信頼リンクが結ばれたような活動的なユーザの集合 \mathcal{V} に焦点を当てる。ここでは、期間 I_0 はさほど長くはない (たとえば $\Delta t_0 = |I_1|$ が 5, 6 程度) と仮定する。 $\mathcal{V} = \{v_1, \dots, v_N\}$ とし、 \mathcal{V} において現在までに生成された信頼リンクを表す $N \times N$ 行列を $G = (G_{i,j})_{i,j=1,\dots,N}$ とおく。ここに、現在までにユーザ v_i からユーザ v_j への信頼リンクが生成されているならば $G_{i,j} = 1$ 、さもなければ $G_{i,j} = 0$ である。本論文では、予測期間 I_1 が短い (たとえば Δt_1 が 2, 3 程度) という仮定の下に、次の Δt_1 か月内に \mathcal{V} において生成される信頼リンクを予測する問題を扱う。ここに、 $\Delta t_0 = |I_0|$ と $\Delta t_1 = |I_1|$ の値は、ユーザの行動パターンは短い期間 (たとえば数か月) においては大きく変化しないという事実が考慮されていることに注意しておく。特に、我々はこの予測問題を行列 G を分解する NMF アプローチにより解くことを考える。

信頼リンク予測問題では、リンク情報の他にも考慮すべき情報がある。 \mathcal{V} に属するユーザが現在までにレビュー投稿や評点付けを行ったアイテムの集合を $\mathcal{A} = \{a_1, \dots, a_M\}$ とし、現在までのユーザアクティビティを表す $N \times M$ 行列を $X = (X_{i,\alpha})$ とする。ここに、 $X_{i,\alpha}$ はユーザ v_i がアイテム a_α に対して付けた評点である。ただし、ユーザ v_i がアイテム a_α に対してレビューを投稿

しなかった場合は $X_{i,\alpha} = 0$ とする。さらに、 \mathcal{V} における現在までのユーザアクティビティに対する評価の情報を表す $N \times M$ 行列を $Y = (Y_{i,\alpha})$ とする。ここに、 $Y_{i,\alpha}$ は、ユーザ v_i がアイテム a_α のレビューに対して得た感謝メッセージの数である。 $hTrust$ と $JLCMF$ は行列 G と X のみを利用したのに対し、本論文では行列 G, X および Y を利用する。

2.2 hTrust

NMF アプローチは潜在空間を用いて有効な特徴量表現を求めるものである。 $hTrust$ は単一の潜在空間を用いる。潜在空間の次元を K とする。 $hTrust$ では非負値 $N \times K$ 行列 $U = (U_{i,k})$ と、非負値 $K \times K$ 行列 $H = (H_{k,\ell})$ を導入する。ここに、 $U_{i,k}$ は潜在因子 k に対するユーザ v_i の強さを、 $H_{k,\ell}$ は信頼リンク生成に対する潜在因子 k から潜在因子 ℓ への関係の強さをそれぞれ表す。 U と H の関数

$$\mathcal{F}_0(U, H) = \|G - UHU^T\|^2 + \lambda_U \|U\|^2 + \lambda_H \|H\|^2 + \lambda_X \text{Tr}(U^T S_X U) \quad (1)$$

の最小化を考える。ここに、 λ_U, λ_H および λ_X は正定数(超パラメータ)である。ただし、任意の行列 B に対して、 B^T 、 $\|B\|$ および $\text{Tr}(B)$ はそれぞれ B の転置行列、 B のフロベニウスノルム、および B のトレースを表す。 $S_X = ((S_X)_{i,j})$ は、

$$(S_X)_{i,j} = \bar{\xi}_i \delta_{i,j} - \xi_{i,j} \quad (2)$$

で定義される $N \times N$ 対称行列である。ここに、 $\delta_{i,j}$ はクロネッカーのデルタであり、

$$\xi_{i,j} = \frac{\sum_{\alpha=1}^M X_{i,\alpha} X_{j,\alpha}}{\sqrt{\sum_{\alpha=1}^M X_{i,\alpha}^2} \sqrt{\sum_{\alpha=1}^M X_{j,\alpha}^2}} \quad (3)$$

かつ、

$$\bar{\xi}_i = \sum_{j=1}^N \xi_{i,j} \quad (4)$$

である。ただし、 $\text{Tr}(U^T S_X U)$ は類同性正則化 (*homophily regularization*) [14] と呼ばれる項であり、

$$\text{Tr}(U^T S_X U) = \frac{1}{2} \sum_{i,j=1}^N \xi_{i,j} \sum_{k=1}^K (U_{i,k} - U_{j,k})^2 \quad (5)$$

であることに注意しておく。この項は類似度の高いユーザ同士はより信頼関係を確立しやすいという性質を組み込むために導入されている。ここに、 $\xi_{i,j}$ はアクティビティに関するユーザ v_i と v_j の間の類似度である。 $\xi_{i,j}$ については別の尺度の選択が可能であり、Jaccard 係数や Pearson 相関係数を含む様々な尺度を使うことができる

ことに注意しておく。非負値拘束 $U \geq 0, H \geq 0$ の下での関数 $\mathcal{F}_0(U, H)$ の最小化問題に対して、 $hTrust$ は U と H の反復更新アルゴリズムを提案した(詳細については [14] を参照されたい)。そのアルゴリズムにより得られた最適な U と H をそれぞれ $U^* = (U_{i,k}^*)$ 、 $H^* = (H_{k,\ell}^*)$ とする。我々は $G_{i,j} = 0$ であるペア (i, j) をリンク候補と呼ぶ。 $hTrust$ は $G_{i,j}^* = \sum_{k,\ell=1}^K U_{i,k}^* H_{k,\ell}^* U_{j,\ell}^*$ の値に従ってリンク候補 (i, j) をランキングすることにより、期間 I_1 に生成される信頼リンクを予測する。

2.3 Joint Link-Content Matrix Factorization

$JLCMF$ も単一の潜在空間を利用する。潜在空間の次元を K とする。 $JLCMF$ では非負値 $N \times K$ 行列 $U = (U_{i,k})$ 、非負値 $K \times K$ 行列 $H = (H_{k,\ell})$ 、および非負値 $M \times K$ 行列 $\Phi = (\Phi_{\alpha,k})$ を導入する。ここに、 $\Phi_{\alpha,k}$ はアイテムへのレビュー投稿の観点における、アイテム a_α と潜在因子 k の間の関係の強さを表す。 U, H および Φ の関数

$$\mathcal{F}_1(U, H, \Phi) = \|G - UHU^T\|^2 + \lambda_H \|H\|^2 + \lambda_\Phi \|\Phi\|^2 + \lambda_X \|X - U\Phi^T\|^2 \quad (6)$$

の最小化を考える。ここに、 λ_H, λ_Φ および λ_X は正定数(超パラメータ)である。非負値拘束 $U \geq 0, H \geq 0$ の下での $\mathcal{F}_1(U, H, \Phi)$ の最小化問題に対して、 $hTrust$ [14] と通常の NMF [16] を応用することにより、 U, H および Φ の反復更新アルゴリズムを容易に得ることができる。 $JLCMF$ は Web ページの分類を目的としているが、信頼リンク予測に対しても適用可能である。したがって、本論文では $JLCMF$ を信頼リンク予測に対してリンクとアクティビティの両方を利用した比較の NMF 法とする。

3 提案法

我々は信頼リンク情報、ユーザアクティビティ情報、およびユーザアクティビティに対する評価情報を組み込んだ新たな NMF 法、すなわち、行列 G, X 、および Y を用いた NMF 法を提案する。まず、信頼リンク予測に対する新たな NMF 法を提案する。次に、提案 NMF 法に対する最適化アルゴリズムを導出する。最後に、ユーザの行動分析への提案 NMF 法の応用法を示す。

3.1 信頼リンク予測のための NMF 法

我々はユーザが好む分野の概念と、ユーザが信頼を獲得する分野の概念を区別することを考える。前者の分野を P 分野、後者の分野を T 分野と呼ぶ。 $hTrust$ や $JLCMF$ とは異なり、提案 NMF モデルは 2 つの潜在空間を用いる。一つは P 分野の空間 (PF 空間と呼ぶ) に対応し、も

う一つは T 分野の空間 (TF 空間と呼ぶ) に対応する。さらに, P 分野と T 分野の各因子をそれぞれ潜在 P 因子と潜在 T 因子と呼ぶ。

PF 空間の次元を K とし, TF 空間の次元を L とする。提案 NMF モデルでは非負値 $N \times K$ 行列 $U = (U_{i,k})$, 非負値 $N \times L$ 行列 $W = (W_{i,\ell})$, および非負値 $K \times L$ 行列 $H = (H_{k,\ell})$ を導入する。ここに, $U_{i,k}$ は潜在 P 因子 k に対するユーザ v_i の強度を表し, $W_{i,\ell}$ は潜在 T 因子 ℓ に対するユーザ v_i の強度を表す。さらに, $H_{k,\ell}$ は信頼リンク生成に対する潜在 P 因子 k から潜在 T 因子 ℓ への関係の強さを表す。 U, W および H の関数

$$\mathcal{F}(U, W, H) = \|G - UHW^T\|^2 + \lambda_U \|U\|^2 + \lambda_W \|W\|^2 + \lambda_H \|H\|^2 + \lambda_X \text{Tr}(U^T S_X U) + \lambda_Y \text{Tr}(W^T S_Y W) \quad (7)$$

の最小化を考える。ここに, $\lambda_U, \lambda_W, \lambda_H, \lambda_X$, および λ_Y は正定数 (超パラメータ) である。 $S_X = ((S_X)_{i,j})$ は式 (2), (3) および (4) によって定義される $N \times N$ 対称行列である。 $S_Y = ((S_Y)_{i,j})$ は, $(S_Y)_{i,j} = \bar{\eta}_i \delta_{i,j} - \eta_{i,j}$ で定義される $N \times N$ 対称行列である。ここに,

$$\eta_{i,j} = \frac{\sum_{\alpha=1}^M Y_{i,\alpha} Y_{j,\alpha}}{\sqrt{\sum_{\alpha=1}^M Y_{i,\alpha}^2} \sqrt{\sum_{\alpha=1}^M Y_{j,\alpha}^2}}$$

かつ,

$$\bar{\eta}_i = \sum_{j=1}^N \eta_{i,j}$$

である。ただし,

$$\text{Tr}(W^T S_Y W) = \frac{1}{2} \sum_{i,j=1}^N \eta_{i,j} \sum_{\ell=1}^L (W_{i,\ell} - W_{j,\ell})^2 \quad (8)$$

である。提案法は正則化項 $\text{Tr}(U^T S_X U)$ と $\text{Tr}(W^T S_Y W)$ により, ユーザアクティビティとユーザアクティビティに対する評価情報を組み込む。ここに, $\xi_{i,j}$ はアクティビティ評価に関するユーザ v_i, v_j 間の類似度であり, $\eta_{i,j}$ はアクティビティ評価に関するユーザ v_i, v_j 間の類似度である。 $\xi_{i,j}$ と $\eta_{i,j}$ については他の尺度の選択が可能であることに再度注意しておく。式 (5) と (8) は, アクティビティの類似するユーザは潜在 P 因子の点で類似する表現を持ち, また, アクティビティ評価の類似するユーザは潜在 T 因子の点で類似する表現を持つということの意味する。さらに, 項 $\|U\|^2, \|W\|^2$ および $\|H\|^2$ は過学習を防ぐための平滑正則化として加えられている。

拘束条件 $U \geq 0, W \geq 0$ および $H \geq 0$ の下での $\mathcal{F}(U, W, H)$ の最小化のために, U, W および H の反復更新アルゴリズム (3.2 節参照) を示す。 $U^* = (U_{i,k}^*)$, $W^* = (W_{i,\ell}^*)$ および $H^* = (H_{k,\ell}^*)$ をそれぞれそのアルゴリズムにより得られた U, W および H の最適値とする。提

案法はリンク候補 (i, j) を $G_{i,j}^* = \sum_{k=1}^K \sum_{\ell=1}^L U_{i,k}^* H_{k,\ell}^* W_{j,\ell}^*$ の値に従ってランキングすることにより, 期間 I_1 に生成される信頼リンクを予測する。

3.2 最適化アルゴリズム

非負値 $N \times K$ 行列 $U = (U_{i,k})$, 非負値 $N \times L$ 行列 $W = (W_{i,\ell})$ および非負値 $K \times L$ 行列 $H = (H_{k,\ell})$ に対して, 式 (7) で定義される関数 $\mathcal{F}(U, W, H)$ の最小化問題を考える。本節ではこの最適化問題を解くための U, W および H の反復更新アルゴリズムを提案する。以降, 任意の行列 B に対して, $B_{p,q}$ は B の (p, q) 成分を示すものとする。

\hat{U}, \hat{W} および \hat{H} をそれぞれ U, W および H の現在の推定値とする。以下において U, W および H の更新式を導出する。まず, 3 つの補助関数

$$\begin{aligned} \mathcal{E}_U(U, \hat{U}; W, H) = & \|G\|^2 + \sum_{i=1}^N \sum_{k=1}^K \frac{(\hat{U}HW^TWH^T)_{i,k}}{\hat{U}_{i,k}} U_{i,k}^2 \\ & - 2 \sum_{i=1}^N \sum_{k=1}^K (GWH^T)_{i,k} \hat{U}_{i,k} \left(1 + \log \frac{U_{i,k}}{\hat{U}_{i,k}}\right) + \lambda_U \|U\|^2 \\ & + \lambda_X \sum_{i,j=1}^N \xi_{i,j} \sum_{k=1}^K \left\{ U_{i,k}^2 - \hat{U}_{i,k} \hat{U}_{j,k} \left(1 + \log \frac{U_{i,k} U_{j,k}}{\hat{U}_{i,k} \hat{U}_{j,k}}\right) \right\} \\ & + \lambda_W \|W\|^2 + \lambda_H \|H\|^2 + \lambda_Y \text{Tr}(W^T S_Y W), \quad (9) \end{aligned}$$

$$\begin{aligned} \mathcal{E}_W(W, \hat{W}; U, H) = & \|G\|^2 + \sum_{i=1}^N \sum_{\ell=1}^L \frac{(\hat{W}H^T U^T UH)_{i,\ell}}{\hat{W}_{i,\ell}} W_{i,\ell}^2 \\ & - 2 \sum_{i=1}^N \sum_{\ell=1}^L (G^T UH)_{i,\ell} \hat{W}_{i,\ell} \left(1 + \log \frac{W_{i,\ell}}{\hat{W}_{i,\ell}}\right) + \lambda_W \|W\|^2 \\ & + \lambda_Y \sum_{i,j=1}^N \eta_{i,j} \sum_{\ell=1}^L \left\{ W_{i,\ell}^2 - \hat{W}_{i,\ell} \hat{W}_{j,\ell} \left(1 + \log \frac{W_{i,\ell} W_{j,\ell}}{\hat{W}_{i,\ell} \hat{W}_{j,\ell}}\right) \right\} \\ & + \lambda_U \|U\|^2 + \lambda_H \|H\|^2 + \lambda_X \text{Tr}(U^T S_X U), \quad (10) \end{aligned}$$

$$\begin{aligned} \mathcal{E}_H(H, \hat{H}; U, W) = & \|G\|^2 + \sum_{k=1}^K \sum_{\ell=1}^L \frac{(U^T U \hat{H} W^T W)_{k,\ell}}{\hat{H}_{k,\ell}} H_{k,\ell}^2 + \lambda_H \|H\|^2 \\ & - 2 \sum_{k=1}^K \sum_{\ell=1}^L (U^T G W)_{k,\ell} \hat{H}_{k,\ell} \left(1 + \log \frac{H_{k,\ell}}{\hat{H}_{k,\ell}}\right) + \lambda_U \|U\|^2 \\ & + \lambda_W \|W\|^2 + \lambda_X \text{Tr}(U^T S_X U) + \lambda_Y \text{Tr}(W^T S_Y W). \quad (11) \end{aligned}$$

を定義する。このとき, 以下の不等式と等式が成り立つ

ことは容易に証明される。

$$\begin{aligned}
\mathcal{E}_U(U, \hat{U}; W, H) &\geq \mathcal{F}(U, W, H), \\
\mathcal{E}_W(W, \hat{W}; U, H) &\geq \mathcal{F}(U, W, H), \\
\mathcal{E}_H(H, \hat{H}; U, W) &\geq \mathcal{F}(U, W, H), \\
\mathcal{F}(U, W, H) &= \mathcal{E}_U(U, U; W, H) = \mathcal{E}_W(W, W; U, H) \\
&= \mathcal{E}_H(H, H; U, W) \quad (12)
\end{aligned}$$

したがって、式 (9) と (12) から、 $\mathcal{E}_U(U, \hat{U}; W, H)$ を U について最小化することで、 $U = (U_{i,k})$ に対する以下の更新式が導出される。

$$U_{i,k} = \hat{U}_{i,k} \sqrt{\frac{(GWH^T)_{i,k} + \lambda_X \sum_{j=1}^N \xi_{i,j} \hat{U}_{j,k}}{(\hat{U}HW^TWH^T)_{i,k} + \lambda_U \hat{U}_{i,k} + \lambda_X \xi_i \hat{U}_{i,k}}} \quad (13)$$

また、式 (10) と (12) から、 $\mathcal{E}_W(W, \hat{W}; U, H)$ を W について最小化することで、 $W = (W_{i,\ell})$ に対する以下の更新式が導出される。

$$W_{i,\ell} = \hat{W}_{i,\ell} \sqrt{\frac{(G^T UH)_{i,\ell} + \lambda_Y \sum_{j=1}^N \eta_{i,j} \hat{W}_{j,\ell}}{(\hat{W}H^T U^T UH)_{i,\ell} + \lambda_W \hat{W}_{i,\ell} + \lambda_Y \eta_i \hat{W}_{i,\ell}}} \quad (14)$$

同様に、式 (11) と (12) から、 $\mathcal{E}_H(H, \hat{H}; U, W)$ を H について最小化することで、 $H = (H_{k,\ell})$ に対する以下の更新式が導出される。

$$H_{k,\ell} = \hat{H}_{k,\ell} \sqrt{\frac{(U^T GW)_{k,\ell}}{(U^T U \hat{H} W^T W)_{k,\ell} + \lambda_H \hat{H}_{k,\ell}}} \quad (15)$$

次に、更新式 (13), (14), (15) を繰り返し用いて導かれる、 U, W および H の反復更新アルゴリズムを説明する。任意の非負整数 t に対して、 U, W および H の t 回目の更新値をそれぞれ U_t, W_t および H_t とする。 U_{t+1} を

$$U_{t+1} = \arg \min_U \mathcal{E}_U(U, U_t; W_t, H_t)$$

により定義すると、これは式 (13) から得られる。また、 W_{t+1} を

$$W_{t+1} = \arg \min_W \mathcal{E}_W(W, W_t; U_{t+1}, H_t)$$

と定義すると、これは式 (14) から得られる。同様に、 H_{t+1} を

$$H_{t+1} = \arg \min_H \mathcal{E}_H(H, H_t; U_{t+1}, W_{t+1})$$

と定義すると、これは式 (15) から得られる。このとき、式 (12) より、

$$\begin{aligned}
\mathcal{F}(U_t, W_t, H_t) &= \mathcal{E}_U(U_t, U_t; W_t, H_t) \geq \mathcal{E}_U(U_{t+1}, U_t; W_t, H_t) \\
&\geq \mathcal{F}(U_{t+1}, W_t, H_t) = \mathcal{E}_W(W_t, W_t; U_{t+1}, H_t) \\
&\geq \mathcal{E}_W(W_{t+1}, W_t; U_{t+1}, H_t) \geq \mathcal{F}(U_{t+1}, W_{t+1}, H_t) \\
&= \mathcal{E}_H(H_t, H_t; U_{t+1}, W_{t+1}) \geq \mathcal{E}_H(H_{t+1}, H_t; U_{t+1}, W_{t+1}) \\
&\geq \mathcal{F}(U_{t+1}, W_{t+1}, H_{t+1})
\end{aligned}$$

を得る。したがって、目的関数 $\mathcal{F}(U, W, H)$ はこの反復更新アルゴリズムにより単調に減少して行く。ゆえにこのアルゴリズムは収束する。

3.3 行動分析への応用

提案アルゴリズムにより得られた U, W および H , すなわち $U^* = (U_{i,k}^*)$, $W^* = (W_{i,\ell}^*)$ および $H^* = (H_{k,\ell}^*)$ を用いて、我々は対象とするソーシャルメディアサイトにおけるユーザの行動分析を行う。潜在 P 因子 (すなわち、P 分野) および潜在 T 因子 (すなわち、T 分野) を調べるために、非負値 $M \times K$ 行列 $\Phi = (\Phi_{\alpha,k})$ および非負値 $M \times L$ 行列 $\Psi = (\Psi_{\alpha,\ell})$ を導入する。ただし、 $\Psi_{\alpha,k}$ はアイテムに対するレビュー投稿の観点からのアイテム a_α と潜在 P 因子 k の間の関係の強さを表し、また、 $\Psi_{\alpha,\ell}$ はアイテムのレビューに対して感謝メッセージを受け取るという観点からのアイテム a_α と潜在 T 因子 ℓ の間の関係の強さを表す。ここで、以下の関数

$$\mathcal{G}_X(\Phi) = \|X - U^* \Phi^T\|^2 + \lambda_\Phi \|\Phi\|^2$$

および、

$$\mathcal{G}_Y(\Psi) = \|Y - W^* \Psi^T\|^2 + \lambda_\Psi \|\Psi\|^2$$

の最小化を考える。ただし、 λ_Φ と λ_Ψ は正定数 (超パラメータ) である。これらの最小化問題に対して、3.2 節と同様の方法で Φ と Ψ の反復更新アルゴリズムを導出することができる。これらのアルゴリズムにより得られた Φ と Ψ の最適値をそれぞれ $\Phi = (\Phi_{\alpha,k})$ と $\Psi = (\Psi_{\alpha,\ell})$ とする。 Φ^* と Ψ^* を用いて、我々はアイテムに関する潜在 P 因子と潜在 T 因子の解釈を試みる。そして、それらの解釈に基づき、信頼リンク生成の観点からサイトにおけるユーザの行動を分析する。

4 評価実験

アイテムレビューサイト Epinions の実データを用いて提案法を評価する。まず、データセットの統計解析から始め、その後、2.1 節に示した信頼リンク予測問題の解法としての提案法の評価を行う。次に、提案法を応用することで、信頼リンク生成の観点からユーザ行動の性質を分析する。

4.1 ソーシャルメディアデータ

我々はアイテムレビューサイト Epinions¹ の実データを収集した。ここに、Epinions は製品レビューと消費者

¹<http://www.epinions.com/>

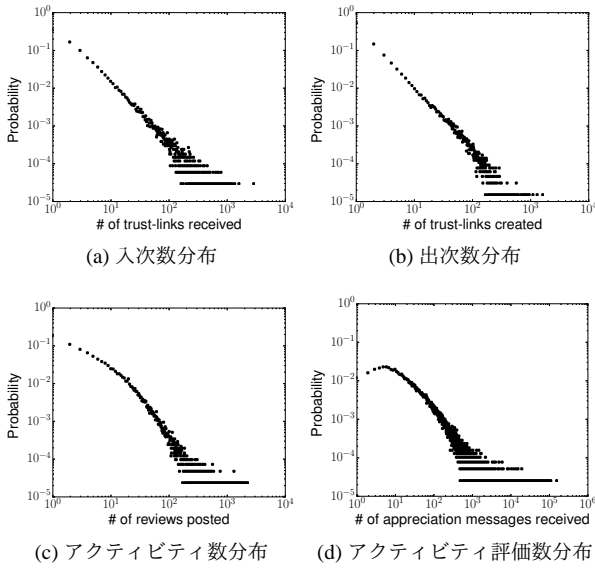


図 1: Epinions データの基本統計解析.

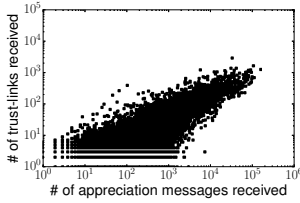


図 2: アクティビティ評価数と入次数の相関.

報告のソーシャルメディアサイトである。Epinionsにおいて、ユーザは他のユーザへの信頼リンクの生成のみならず、与えられたアイテム群のアイテムに対してレビュー投稿や評点付けを行うことができる。さらに、あるユーザのアイテムへのレビューに対して、別のユーザはそれを有益であると思ったときに感謝メッセージを投稿することができる。Epinionsに関して、我々は2012年10月に最人気ユーザとして紹介されたユーザを起点として、新たなユーザが出現しなくなるまで信頼リンクを幅優先探索し、信頼リンク、レビューと評点、および感謝メッセージのデータを収集した。収集したデータは、ユーザ数 64,268 人、信頼リンク数 509,293 本、アイテム数 268,891 個、レビュー数 809,517 件および感謝メッセージ数 18,960,792 件であった。

Epinions データに対して、まず、基本的な統計的性質を分析するために、入次数分布（ユーザが受け取った信頼リンク数の割合）、出次数分布（ユーザが生成した信頼リンクの割合）、アクティビティ分布（ユーザが投稿したレビュー数の割合）およびアクティビティ評価分布（ユーザが受け取った感謝メッセージ数の割合）を解析した。図 1 にその結果を示す。いずれの分布も後部がべき乗則に従うことが観測される。これらの結果は、Epinions データがオンライン世界におけるソーシャルデータの典型的な性質を満足することを示している。次に、アクティビティ評価数と入次数の間の相関を調査した。図 2 にその結果を示す。ユーザが受け取った感謝メッセージ数とユーザが受け取った信頼リンク数との間には正の相関があることが分かる。これは、信頼リンク予測の NMF 法がアクティビティ評価情報を組み込むことにより改善される可能性を示唆している。

4.2 実験設定

Epinions データを用いて、信頼リンク予測問題 (2.1 節参照) の解法である提案法の性能を評価した。Epinions データに対して、4 つのデータセット D_1, D_2, D_3 および D_4 を以下のように構築した。まず、 $n_1 = n_2 = 1, \Delta t_0 = 6$, および $\Delta t_1 = 3$ (2.1 節参照) と設定した。評価期間 I_1 は、 D_1 に対しては 1 月から 3 月、 D_2 に対しては 4 月から 6 月、 D_3 に対しては 7 月から 9 月、 D_4 に対しては 10 月から 12 月とした。2006 年には一定数以上の信頼リンクが常に生成されていたことからその年に生成された信頼リンクの予測を試みた。ここに、たとえば、データセット D_1 に対する I_1 は 2006 年の 1 月から 3 月であり、観測期間 I_0 は 2005 年の 7 月から 12 月である。表 1 は各データセットの基本統計である。

表 1: データセットの基本統計.

	D_1	D_2	D_3	D_4
ユーザ数, N	771	722	734	727
アイテム数, M	56,642	57,886	59,522	61,327
観測信頼リンク数	27,154	25,382	26,096	26,581
アクティビティ数	83,786	83,933	86,396	88,980
アクティビティ評価数	1,074,042	1,039,505	1,127,326	1,170,950
I_1 に生成された信頼リンク数	1,670	1,308	1,441	1,684

信頼リンク予測問題に対して、我々は提案法と $hTrust$ (式 (1) 参照) および $JLCMF$ (式 (6) 参照) を比較する。また、 $JLCMF$ を拡張することで提案法に類似する手法を定義し、提案法との比較を行う。ここに、非負値 $N \times K$ 行列 U , 非負値 $N \times L$ 行列 W , 非負値 $K \times L$ 行列 H , 非負値 $M \times K$ 行列 Φ および非負値 $M \times L$ 行列 Ψ に対して、以下の関数

$$\mathcal{F}_2(U, W, H, \Phi, \Psi) = \|G - UHW^T\|^2 + \lambda_U \|U\|^2 + \lambda_W \|W\|^2 + \lambda_H \|H\|^2 + \lambda_X \|X - U\Phi^T\|^2 + \lambda_Y \|Y - W\Psi^T\|^2 + \lambda_\Phi \|\Phi\|^2 + \lambda_\Psi \|\Psi\|^2$$

の最小化を考える。ただし、 $\lambda_U, \lambda_W, \lambda_H, \lambda_X, \lambda_Y, \lambda_\Phi$ および λ_Ψ は正定数 (超パラメータ) である。この最小化問題に対して、提案法 (3.2 節参照) と同様の導出により U, W, H, Φ および Ψ の反復更新アルゴリズムが容易に得られる。我々はこの手法を $JLCMF2$ と呼ぶ。

各手法における超パラメータは交差検証により決定することが可能であるが、本論文では [14] に従って単純に $\lambda_U = \lambda_W = \lambda_H = \lambda_\Phi = \lambda_\Psi = 0.01$, $K = L = 10$, $\lambda_X = \lambda_Y = 10$ と設定した²。

4.3 信頼リンク予測

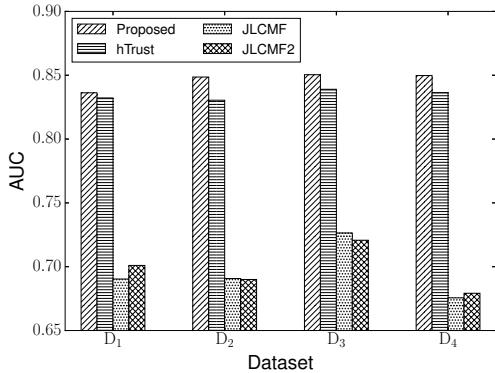


図 3: 信頼リンク予測性能の比較。

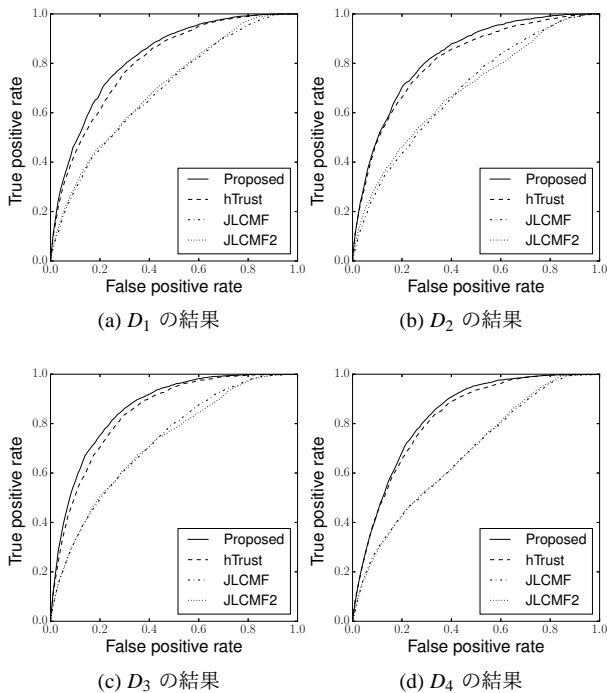


図 4: ROC 曲線。

信頼リンク予測問題 (2.1 節参照) に対して、提案法と *hTrust*, *JLCMF* および *JLCMF2* を比較する。実験では、ROC 曲線下の面積 (AUC) で予測性能を測定した。図 3 にその結果を示す。図 4 は ROC 曲線の結果である。提

²超パラメータの影響を詳細に調査することは今後の課題である。

案法が最も高性能であり、*hTrust* がそれに続いていることが観察される。性能差はデータ依存であるが、*JLCMF* と *JLCMF2* はこれら 2 つの手法を常に下回っていた。これらの結果はアクティビティ評価情報を組み込むことの重要性を示しており、2 種類の潜在因子 (すなわち、2 つの潜在空間) である潜在 P 因子および潜在 T 因子を適切に融合した提案法の有効性を実証している。

4.4 行動分析

提案法を応用し、*Epinions* での信頼リンク生成に対するユーザ行動の分析を行う。ここでは、データセット D_3 に対する分析結果のみを報告する。

図 5 に $H^* = (H_{k,\ell}^*)$ の可視化結果を示す。 k 行 ℓ 列の明るさは、潜在 P 因子 k から潜在 T 因子 ℓ への関係の強さ $H_{k,\ell}^*$ の値を表している。 $H_{k,\ell}^*$ は $(k, \ell) = (6, 6)$ において最大であった。また、 $\Phi^* = (\Phi_{\alpha,k}^*)$ と $\Psi^* = (\Psi_{\alpha,\ell}^*)$ を用いて、潜在 P 因子 $\{k\}$ と潜在 T 因子 $\{\ell\}$ がどのようなものを調べた。その結果、潜在 P 因子と潜在 T 因子の間には明らかな違いが見られたことに注意しておく。

表 2 と表 3 にそれぞれ $\Phi_{\alpha,6}^*$ と $\Psi_{\alpha,6}^*$ に関する上位 4 アイテム $\{a_\alpha\}$ を示す。これらは潜在 P 因子 $k = 6$ と潜在 T 因子 $\ell = 6$ を例示していることに注意しておく。したがって、*Epinions* のデータセット D_3 では、テレビゲームと SF 映画を好むユーザから SF・ファンタジーで信頼を獲得するユーザへと信頼リンクが結ばれる傾向があることが分かる。

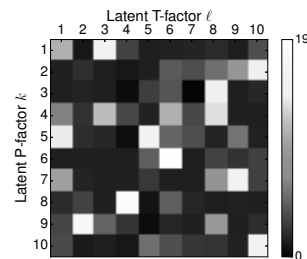


図 5: D_3 における $H^* = (H_{kl}^*)$ の可視化。

5 結論

本論文では、アイテムレビューサイト (ソーシャルメディアサイト) に対して、直近の過去において活動的な

表 2: D_3 における潜在 P 因子の例 ($k = 6$)。

Item	Category
Nintendo Game Cube White Console	Video Game Consoles
Sony PlayStation 2 Slimline Console	Video Game Consoles
Star Wars Episode III: Revenge of the Sith	Movies
Sega Dreamcast Grey Console	Video Game Consoles

表 3: D_3 における潜在 T 因子の例 ($\ell = 6$).

Item	Category
Lord of the Rings: The Return of the King	Movies
Star Wars Episode III: Revenge of the Sith	Movies
The Incredibles by Pixar & Disney	Movies
Lord of the Rings: The Fellowship of the Ring	Movies

ユーザを対象に、直近の未来に生成される信頼リンクを予測する問題を扱った。近年 NMF 法を用いた $hTrust$ の有効性が示されたことから、我々はこの問題に対して NMF アプローチを採用した。 $hTrust$ は信頼リンク情報とユーザアクティビティ情報を用いる。ここに、ユーザアクティビティはアイテムに対してレビュー投稿や評点付けを行うことを意味する。多くのアイテムレビューサイトでは、有益であったレビューに感謝メッセージを投稿することができる。つまり、ユーザアクティビティ評価情報をサイト内にて観測することができる。信頼リンク予測に対する NMF 法の性能の向上を目指して、我々は信頼リンク情報とアクティビティ情報に加えてアクティビティ評価情報を組み込んだ新たな NMF 法を提案した。そして、アイテムレビューサイト Epinions の実データを用いて提案法の評価を行った。

まず、我々は受け取った感謝メッセージ数と受け取った信頼リンク数の間に相関があることを示し、アクティビティ評価情報を組み込むことが有効である可能性を示した。次に、信頼リンク予測問題において提案法が $hTrust$ やそれと類似した手法 $JLCMF$ および $JLCMF2$ と比べて高性能であることを示した。これらの結果はアクティビティ評価情報を組み込むことの重要性を示しており、2 種類の潜在因子 (すなわち、2 つの潜在空間) である潜在 P 因子および潜在 T 因子を適切に融合した提案法の有効性を実証している。さらに、提案法をアイテムレビューサイトにおけるユーザの行動分析に応用し、信頼リンク生成の観点から Epinions のいくつかの特徴的な性質を明らかにした。

謝辞

本研究は JSPS 科研費 26330352 の助成を受けたものである。

参考文献

- [1] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the spread of influence through a social network,” in *ACM SIGKDD*, 2003, pp. 137–146.
- [2] M. Kimura, K. Saito, and H. Motoda, “Blocking links to minimize contamination spread in a social network,” *ACM TKDD*, vol. 3, pp. 9:1–9:23, 2009.
- [3] M. Kimura, K. Saito, R. Nakano, and H. Motoda, “Extracting influential nodes on a social network for information

diffusion,” *Data Mining and Knowledge Discovery*, vol. 20, pp. 70–97, 2010.

- [4] K. Saito, M. Kimura, K. Ohara, and H. Motoda, “Selecting information diffusion models over social networks for behavioral analysis,” in *ECML-PKDD*, ser. LNAI, vol. 6323. Springer, 2010, pp. 180–195.
- [5] W. Chen, L. Lakshmanan, and C. Castillo, “Information and influence propagation in social networks,” *Synthesis Lectures on Data Management*, vol. 5, pp. 1–177, 2013.
- [6] M. Kimura, K. Saito, K. Ohara, and H. Motoda, “Learning to predict opinion share and detect anti-majority opinionists in social networks,” *Journal of Intelligent Information Systems*, vol. 41, pp. 5–37, 2013.
- [7] K. Saito, M. Kimura, K. Ohara, and H. Motoda, “Detecting changes in information diffusion patterns over social networks,” *ACM TIST*, vol. 4, pp. 55:1–55:23, 2013.
- [8] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, “Propagation of trust and distrust,” in *ACM WWW*, 2004, pp. 403–412.
- [9] D. Liben-Nowell and J. Kleinberg, “The link-prediction problem for social networks,” *Journal of the American Society for Information Science and Technology*, vol. 58, pp. 1019–1031, 2007.
- [10] H. Liu, E. Lim, H. Lauw, M. Le, A. Sun, J. Srivastava, and Y. Kim, “Predicting trusts among users of online communities: an opinion case study,” in *ACM EC*, 2008, pp. 310–319.
- [11] H. Mannila and E. Terzi, “Finding links and initiators: a graph-reconstruction problem,” in *SIAM SDM*, 2009, pp. 1207–1217.
- [12] V. Nguyen, E. Lim, J. Jiang, and A. Sun, “To trust or not to trust? predicting online trusts using trust antecedent framework,” in *IEEE ICDM*, 2009, pp. 896–901.
- [13] M. Gomez-Rodriguez, J. Leskovec, and A. Krause, “Inferring networks of diffusion and influence,” in *ACM SIGKDD*, 2010, pp. 1019–1028.
- [14] J. Tang, H. Gao, X. Hu, and H. Liu, “Exploiting homophily effect for trust prediction,” in *ACM WSDM*, 2013, pp. 53–62.
- [15] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri, “Feedback effects between similarity and social influence in online communities,” in *ACM SIGKDD*, 2008, pp. 160–168.
- [16] D. Lee and H. Seung, “Algorithms for non-negative matrix factorization,” in *NIPS*, 2001, pp. 556–562.
- [17] S. Zhu, K. Yu, Y. Chi, and Y. Gong, “Combining content and link for classification using matrix factorization,” in *ACM SIGIR*, 2007, pp. 487–494.

セルフサービスシステムの実現に向けた回答データ作成手法 Method for collecting answer data for constructing self-service system

竹内広宜^{1*} 増田聡¹ 宮本晃太郎¹
Hironori Takeuchi¹ Satoshi Masuda¹ Kohtaroh Miyamoto¹

¹ 日本アイ・ビー・エム株式会社 東京基礎研究所
¹ IBM Research - Tokyo

Abstract: By using the question answering technology, we can develop a self-service system for various business. When constructing a self-service system, we need to collect pairs of questions and answers. For the answer data, it is required that collected data covers whole target business. In this paper, we consider the method to collect sufficient answer data by modeling the user scene and the target service and evaluate its effectiveness in a real project.

1 はじめに

企業は競争力を維持するために、ビジネス環境の急激な変化にあわせ、先革新的なビジネスを作り出すことを重要視している。このような革新的なビジネスはイノベーションと呼ばれ、その創出過程に関する研究が行われている。例えば、新サービスの創出（サービスデザイン）を行うためにデザイン思考 [4] と呼ばれる試みが研究され、実践されつつある。

一方、クラウド技術の発展に伴い、アナリティクスをはじめとした様々な技術が API やデモとしてクラウド上で公開されている。これにより、先進技術を活用した新しいビジネスアイデアもデザイン思考を通して創出されるようになってきている。先進技術を活用した新しいアイデアのシステム化にあたっては、必要となる知識データといった前提条件を非技術者も含め整理し、実現に向けた検討を行う必要がある。この時、アイデアの実現に向け必要な前提条件を洗い出す手法が提案されており [6]、これにより、システム化に向けた計画が立てられるようになっている。しかしながら、この段階では、知識データなど前提条件がわかるものの、それを満たすためにどのような作業が必要かはわからない。実現する内容に基づき、別途、作業規模の見積もりや手順の明確化が必要となる。

自然言語処理や機械学習といった技術を活用したシステムを構築する場合は、対象分野向けの辞書や学習データが必要となる。ここで、どのような辞書や学習データといったデータ資源をどのくらい作成すればよ

いかは、対象となる目的によって異なる。その結果、自然言語処理や機械学習の技術を適用するシステム構築プロジェクトでは以下の課題が生じる可能性がある。

- データ資源の作成の工数や期間といったプロジェクトの規模見積もりが不十分となる。
- 作成するデータ資源が対象ビジネスをカバーしているかどうかを関係者で共通理解し合意することができない。
- テスト範囲が不明確となる。また、テスト段階で期待通りの結果が得られない場合、作成したデータ資源が原因かどうかの切り分けが困難となる。

このような技術を適用するには、必要となるデータ資源が十分であるかを検証することが非常に重要となる。

本研究では、自然言語処理や機械学習の技術を適用するシステムの構築に必要なデータ資源（リソース）の作成において、対象分野に対して網羅性を持ったリソースを作成する手法を考える。構築するシステムは、あるビジネスについて利用者に、要求に合致した情報を提供するものと考え、本研究では、対象ビジネス（サービス）および利用者（ユーザー）がシステムを使う状況をモデル化し、リソース作成に活用する手法を提案する。具体的なシステムとして、質問応答技術を利用し顧客からの問い合わせに自動対応するセルフサービスシステムを想定し、実際のシステム構築の際に作成したリソースを元に提案した手法の有効性を検討する。

本論文の構成は以下の通りである。まず 2 節で本研究で対象とするシステムについて述べ 3 節で、そのシステムを構築する上での課題とデータ資源に求められる性質について述べる。その後、4 節で対象サービスと

*連絡先：日本アイ・ビー・エム株式会社 東京基礎研究所
〒103-8510 東京都中央区日本橋箱崎町 19-21
E-mail: HIRONORI@jp.ibm.com

利用するユーザーの状況のモデル化に基づくデータ資源の定義手法について述べる。5節で提案手法の適用したケーススタディとその考察結果について述べ、そして、6節でまとめを行う。

2 問い合わせ対応セルフサービスシステム

サービス（商品も含む）に関する情報は社内だけでなく社外にも Web ページなどの形式で大量に公開されている。また、サービスによっては提供者側からの情報だけでなく口コミなど利用者側から提供される情報も存在する。

サービスのユーザーは、自身がおかれている状況に応じて何らかの目的を持っている。しかし、その目的を達成するために現在持っている知識では不十分であると感じる状態がある。この時、ユーザーが抱くものを情報要求と呼び、ユーザーの情報要求を満たす技術の総称を情報アクセス技術と呼んでいる [8]。

情報アクセス技術を利用したシステムとして、情報検索システムは企業内で広く利用されている。また、近年は、ユーザーからの質問に対して自動的に回答する技術として質問応答 (Question Answering:QA) 技術が様々な分野で実用化されつつある。この QA 技術の活用例として、顧客からの問い合わせに自動応答するシステムが考えられる。問い合わせ対応システムは、ある対象領域について、エンドユーザーからの問い合わせや要求に、その領域の専門家の代役として対応するシステムである。顧客対応には、コールセンターでのテクニカルサポートをはじめ様々な問い合わせ業務が該当し、利用シーンも、購入や利用前の潜在顧客に対応するものや、サービスを利用している顧客にアフターサービスとして対応するものなど多岐にわたる。このような顧客からの問い合わせに対しては、待たせることなく、的確な情報を提供することが顧客満足度向上において必要となっている。一方、迅速に的確な情報を提供するためには多数の専門家を 24 時間、コールセンターなどに配置するという課題もある。そのため、顧客とのチャネルのひとつに、顧客からの問い合わせに自動応答するシステムを置くことで、顧客満足度を維持しながら専門家を配置するコストを削減することが期待できる。本研究では、QA 技術を用いてこのような応対自動応答するシステムを対象に考え、以降、問い合わせ対応セルフサービスシステムと呼ぶ。

QA 技術は主に以下の構成要素からなる。

- 質問解析: 入力 of 自然文を解析し、単語や意味情報などを抽出する

- 回答候補検索: 抽出した単語や意味情報に適合する回答を検索する
- 回答候補評価: 検索した得られた回答候補に適合度スコアを付与する

問い合わせ対応セルフサービスシステムでは、入力インターフェースから受け取った自然文テキストを QA 技術コンポーネントに受け渡す。そして、得られたスコア付きの回答候補から、出力インターフェースに応じて必要個数の回答を選択し表示する。

近年、上記で述べた質問応答技術の 3 つの機能をコンポーネント化したものが API として利用できるようになっており、簡単にアプリケーションを作ることが可能となっている。QA 技術コンポーネントについて様々な実装が考えられるが、本研究では分類器学習をベースとしたものを用いる¹。ここでは、回答データ (回答文) $A = \{a_1, a_2, \dots, a_n\}$ を定義し、それぞれの回答文 a_i について想定質問 $Q_i = \{q_{i1}, q_{i2}, \dots, q_{im}\}$ を準備する。ここで n は回答の範囲をあらわし、対象分野に依存する。 m は回答文を得るための想定質問のバリエーションに相当する (例としては $m = 10$)。こうして作られた A, Q_1, \dots, Q_n を学習データとして QA 技術コンポーネント内の機械学習に適用し、問い合わせ対応セルフサービスシステムとして利用可能な状態にする。

3 問い合わせ対応セルフサービスシステムの構築で求められる品質特性とデータ資源に求められる性質

情報アクセス技術の評価は、通常、テストコレクションと呼ばれる、情報要求の適用先となる対象文書データと情報要求 (検索要求) からなるデータを準備する。そして、情報要求に対する結果を比較し、技術の有効性を評価する [5][8]。QA 技術を利用したシステムの評価では、出力インターフェースで表示する回答候補の中にユーザーが必要とする回答が含まれることが重要となる。そこでシステムの出力数 k とした場合、評価用の質問を学習データと独立して準備し、それぞれの質問に対して上位 k 番目までに正解が含まれるかどうかを示す k 位正答率などで性能を測ることが行われる。

しかしながら、上記で述べた評価だけでは問い合わせ対応セルフサービスの実用システムの構築上、いくつかの課題が生じる。ある評価用質問に対して、システムが回答できなかった場合、それが機械学習の結果によるものなのか、そもそも望まれる回答が学習デー

¹例: <https://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/doc/nl-classifier/index.shtml>

タに入っていない、という2つの原因が考えられる。これから、システムを構築する上では、学習に用いたデータが以下の2点について十分であったかの検討が必要となることがわかる。

1. 定義した回答（回答文）データ A は対象サービス全体を網羅しているか
2. 各回答に対する m 個の想定質問のバリエーションは網羅的に作成されているか

本研究では、これらを問い合わせ対応セルフサービスシステムにおいて学習データが満たすべき特性と考える。

ISO9126 ではソフトウェアシステムの品質特性として機能性、信頼性、使用性、効率性、保守性、移植性の6観点が定義されている [1]。このうち、機能性は合目的性（適切性）、正確性、相互運用性、機密性、標準適合性に細分化されている。これらのうち、合目的性と正確性の定義は以下の通りである。

- 合目的性 (suitability): システムは目的に合致している機能を提供する
- 正確性 (accuracy): システムは入力に対して常に期待される結果を出す

機能性のうち、本研究で対象とする問い合わせ対応セルフサービスシステムで重要視される特性は、合目的性と正確性と考えられる。問い合わせ対応セルフサービスシステムの合目的性は、対象業務に必要な情報を全て保持し、出力することに相当する。また、正確性は入力（質問文）に対して合致する情報を出力することに相当する。学習データが満たすべき特性のうち1が合目的性に、そして2が正確性に関係する。

QA 技術コンポーネントの正確性については、テスト段階で想定質問のバリエーションを調整することで改善することができる。一方、合目的性については、テスト段階で不具合が見つかった場合、回答データ全体の精査をした上で必要であれば、回答の追加と想定質問の作成を行う必要がある。そのため、要件定義の段階で、対象とするビジネスの知識情報の範囲について関係者で合意を取ることが重要である。また、テスト段階で不具合が見つかった際の修正を容易にするため、定義する回答データが網羅的であるかを体系立てて管理する必要がある。本研究では、問い合わせ対応セルフサービスシステムの品質特性のうち合目的性を向上させるための回答データ作成方法を検討する。次節でその手順について述べる。

4 対象サービスと利用顧客の分析に基づく回答データ作成手法

4.1 質問タイプと顧客の消費行動

QA 技術において、質問の分類は重要であり、さまざまな分類が定義されている。質問者が自身の質問への回答に期待するタイプを質問タイプと呼んでいる [7]。

現在、Web 上で様々な QA サイトが運営されている。QA サイトの質問を分析し、質問タイプを検討する研究が行われている。QA サイトには、事実・定義の質問だけでなく、提案・意見を求める質問が存在する [2][3]。[9] では、事実や定義を問う質問を情報検索型質問、助言や意見を求める質問を社会調査型質問というカテゴリーを定義し、それぞれを細分化している。

本研究で対象とするサービスに関する問い合わせに対応するセルフサービスシステムでは、情報検索型質問だけでなく、助言や推薦などの社会調査型質問が入力として想定される。[9] で定義された質問タイプのうち以下を問い合わせ対応セルフサービスシステムに必要な対象サービスについての知識と考え、知識タイプと呼ぶ。

1. 事実・定義: 名称, 場所, 日付, 大きさなどの事実や説明
2. 方法・手段: あることを行う方法や手段の説明
3. 理由・結果: ある物事の理由や結果についての記述
4. 助言・推薦: ある物事に対する助言や推薦の記述

各知識タイプについてサービスに関する知識を整理したものを知識データと呼ぶ。この知識データ中の各データが問い合わせ対応セルフサービスシステムで使われる回答文に相当する。したがって、対象サービスを分析し、上述の知識データの形で構造化したものが回答データとなる。

知識タイプに対する情報要求は、対象サービスへの顧客の理解度や利用状況と関係がある。例えば、理由に関する知識は、実際にサービスを使って、該当する事象を体験する段階で必要性が上がる。一方、サービスを利用する段階では事実・定義に関する知識への必要性は低くなる。

顧客が商品やサービスを購入するまでの行動段階については以下の3段階が考えられている。

1. 認知段階: 知る
2. 感情段階: 興味を持ち、購入したいと思う
3. 行動段階: 購入する

このサービスの購買に至るまでの3段階と、それぞれの段階で顧客が必要とするサービスの知識の関係性を表1に示す。

表 1: 購買行動の段階と求められるサービス知識の関係性 (H:高 M:中 L:低)

	認知段階	感情段階	行動段階
事実・定義	H	H	M
方法・手段	M	H	H
理由・結果	L	L	H
助言・推薦	H	M	L

問い合わせ対応セルフサービスシステムが、どの段階にいる顧客からのサービスへの問い合わせに対応するかによって、準備する知識データが変わる。本研究では、システムの想定利用者の段階と関係性が中程度から高い関係性を持つ知識データを整理し回答データを作成することを考える。

次に、対象サービスの構造を分析し、回答データのための知識データを作成する手法を述べる。

4.2 対象サービスの分析と回答データの作成

本研究では、対象となるサービスの機能を以下の要素でモデル化する。

- サービス機能：サービスが持つ機能。複数の機能で構成される場合、それらをサブサービス機能と呼ぶ。
- 可変な属性：数値であらわされる属性または顧客によって変更の可能性がある属性。
- 不変な属性：顧客によって変更の可能性がない属性。

これらの要素の関係を図1に示す。

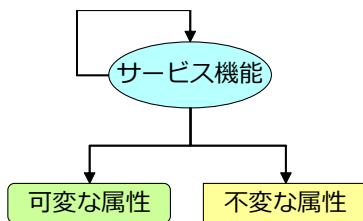


図 1: サービス機能のモデル

このモデルを用いて、以下の手順で対象サービスを分解する。まず、対象サービスについて、可変な属性お

よび不変な属性を列挙する。ここで、可変な属性については、この時点で顧客が変更できるか分析で明確に決定できないものもある。変更できる可能性があると考えられる場合には、可変な属性として列挙する。次に、対象サービスを構成するサブサービス機能を列挙する。そして、それぞれのサブサービス機能に親サービス機能が持つ属性を一時的に継承させる。そして、継承されない属性があれば取り除く。属性を列挙するステップにもどり、サブサービス機能がなくなるまで繰り返す。

上記手順に基づいて、インターネットバンキングサービスを分析した結果を図2に示す。

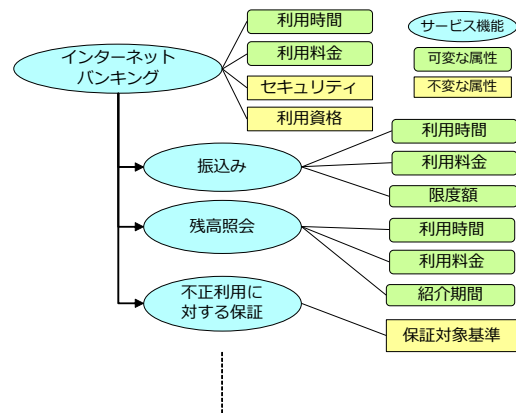


図 2: サービス機能の分解に基づく分析結果

次に、分析によって得られたサービスの構成要素に対して必要となる知識タイプを整理する。構成要素と必要となる知識タイプの関係を表2に示す。

表 2: サービスの構成要素と必要となる知識タイプの関係 (Y:必要あり N:必要なし)

	事実・定義	方法・手段	理由・結果	助言・推薦
サービス機能	Y	Y	Y	Y
可変な属性	Y	Y	Y	N
不変な属性	Y	N	N	N

表1で示したように、システムが対象とする顧客の状態（消費者行動の段階）に応じて必要となる知識タイプが決まる。そして、表2を元に、必要な知識タイプに当てはまる知識を割り当てて整理する。この段階で、可変な属性について、方法・手段についての知識データが作成できない場合、その属性を不変な属性に変更する。例えば、認知段階および感情段階の顧客を対象とした場合、サービス機能については、「サービス

表 3: 回答データに用いられる知識データの整理例

知識トピック	事実・定義	方法・手段	助言・推薦
インターネットバンキング	インターネットバンキングはPCやスマホから銀行取引をするサービスです	インターネットバンキングは口座開設時にお申し込みできます	インターネットバンキングは店舗に行かなくても取引でき大変便利です
残高照会の照会期間	PCやスマホから取引記録を見るサービスでの対象期間です	設定画面で過去2年までを設定することができます	N/A
不正利用への補償の対象基準	不正利用が明らかになったお客様が対象となります	N/A	N/A

の説明(事実・定義)、「サービスの使い方(方法・手段)」、「サービスの良さや訴求店(助言・推薦)」の知識タイプについて当該サービスに関する知識記述を知識データとして整理する。こうして整理された知識データは、顧客からの問い合わせへ対応する問い合わせ対応セルフサービスシステムにおいてシステムが出力する回答文となる。

認知段階および感情段階の顧客を対象としたシステムについて、図2で得られたサービス機能、可変な属性、不変な属性について作成した知識データの例を表3に示す。

こうした作成された回答データを元に、QAシステム用の学習データを作成する。各回答文に対して、想定される質問文を列挙する。例えば、「不正利用が明らかになったお客様が対象となります」というインターネットバンキングの不正利用への補償の対象基準に関する事実・定義の知識タイプを持つ回答(知識)に対して、「不正利用の補償はどういった場合ですか?」、「不正利用の補償となる対象は誰ですか?」、「不正に使われたら誰でも補償されるのですか?」といった想定質問を付与して学習データとする。

5 ケーススタディと考察

銀行サービスにおける問い合わせ対応セルフサービスシステムに対して作成した回答データを用いて、提案手法の検証を行った。銀行サービスには、口座開設、口座取引、インターネットバンキング、外国送金、投資信託など様々な商品・サービスがある。これらの商品・サービスに関して、認知・感情段階の顧客からの問い合わせに対応するセルフサービスシステムを想定し、(個人向け)外国送金サービスを対象に提案手法の適用を行った。その結果、サービス機能が5個、可変な属性が0個、不変な属性が14個同定された。そして、表2に基づき知識データを整理して29個の回答文からなる回答データが作成された。

提案手法の適用と並行し、作成手順やガイドの提示なく、人手による回答データの作成を行った。作成においては、「外国送金の必要性を感じている顧客からの問い合わせ対応システムにおいて、回答すべき情報を列挙する」という指示のもと、作成者は、銀行のWebページの説明などから独自に回答データの作成した。その結果、独自作成では、14個の回答文からなる回答データが作成された。提案手法により、より多くのサービス機能および属性が同定できていることがわかる。

人手による独自作成では、事実・定義に関するサービス知識のみが作成された。また、「組戻し」「被仕向送金の送金到着案内」といったサービス機能とその属性を見落としていた。また、同定できたサービス機能についても、仕向送金における禁止取引先といった属性を見落としていた。一方、提案手法による見落としは、仕向送金における取引制限金額という属性に関するものだけであった。提案手法および独自作成による実践者はそれぞれ1名だけであるため分析者の経験や対象サービスへの理解の差の影響が考えられる。より正確な評価には、複数人による比較実験が必要であるが、本手法により、より網羅的な回答データが作成できることが期待できると考えられる。

提案手法を用いて、要件定義および設計の初期段階で、問い合わせ対応セルフサービスが対象とするサービスに関する知識を構造化して整理し、回答データを共有することで関係者で共通の理解ができるとともに、その網羅性や十分性についての確認と合意ができる。その結果、テスト段階において、回答データの不備が原因による手戻りについて、その発生数を抑えることが期待できる。また、回答データ作成後、各回答文に対する複数の想定質問文を学習データとして準備する必要があるが、その際の作業規模を見積もることも容易となると期待できる。

対象範囲が広いサービスの問い合わせ対応セルフサービスを考えた場合、対象サービスを本研究したモデルを用いて人手で構造分析することはコストがかかるた

め何らかの支援技術が必要となる。例えば、業務マニュアルといった文書を自然言語解析技術により分析し、サービスの構成要素や属性を示す語や表現を抽出することなどが考えられる。また、対象サービスによっては、FAQ サイトなど事前にデータがあることもある。その場合、システム構築の際、そのデータを用いることで回答データの構築を効果的に進められることができるが、FAQ データが問い合わせ対応セルフサービスシステムの回答データとして十分かどうかの評価が必要となる。また、不足部分がある場合は、業務知識全体のどの部分を補足する必要があるを見極める支援が有用であると考えられる。これらの点が今後の課題である。

6 まとめ

本研究では、QA 技術を利用した問い合わせ対応セルフサービスシステムの構築における回答データの作成について検討した。実用システムの構築においては、システムの合目的性を満たすために、回答データの網羅性が必要であることを示した。そして、対象ビジネスを機能および属性に構造分解し、システムを利用する対象顧客の状況を元に必要な知識タイプに展開し、知識データとして整理し回答データとして利用する手法を提案した。実際の銀行サービスを対象としたシステム構築を想定し、提案手法を用いて回答データを作成した。そして、人手による独自作成の結果と比較した結果、提案手法を用いることでより網羅性の高い回答データの作成が期待できることがわかった。対象ビジネスの範囲が広い場合における、サービス機能の構造分解の分析支援や、FAQ データが事前に存在する場合における本提案手法の活用方法などが今後の課題となっている。

参考文献

- [1] S. N. Bhatti. Why quality?: ISO 9126 software quality metrics (functionality) support by UML suite. *SIGSOFT Software Engineering Notes*, 30(2):1–5, 2005.
- [2] S. Kim, J. S. Oh, and S. Oh. Best answer selection criteria in a social q& a site from the user-oriented relevance perspective. In *American Society for Information Science and Technology Annual Meeting*, pp. 1–15, 2007.
- [3] E. M. Rodrigues and N. Milic-Frayling. Socializing or knowledge sharing? Characterizing social intent in community question answering. In *Proceedings*

of the 18th ACM Conference on Information and Knowledge Management (CIKM), pp. 1127–1136, 2009.

- [4] M. Stickdorn and J. Schneider. *This is Service Design Thinking: Basics, Tools, Cases*. Wiley, 2011.
- [5] E. M. Voorhees and D. K. Harman. *TREC: Experiment and Evaluation in Information Retrieval*. The MIT Press, 2005.
- [6] 竹内, 張. 顧客体験記述を元にした新システムの構築検討手法とその実践. 信学技法 KBSE2015-5, pp. 21–26, 2015.
- [7] 渡邊, 島田, 関, 神門, 佐藤. QA コミュニティにおける質問者の期待に基づく質問分類に関する一検討. In *DEIM Forum B5-1*, 2011.
- [8] 酒井. 情報アクセス評価方法論. コロナ社, 2015.
- [9] 栗山, 神門. Q&A サイトにおける質問と回答の分析. 情報処理学会研究報告 2009FI-95-19, pp. 1–8, 2009.

11月10日(木) 発表分の受賞理由

6. SIG-SWO セマンティックウェブとオントロジー研究会

「述語論理に基づく農作業オントロジーの設計と応用」

朱 成敏 (国立情報学研究所), 小出 誠二 (情報・システム研究機構), 武田 英明 (国立情報学研究所), 法隆 大輔 (農研機構), 竹崎 あかね (農研機構), 吉田 智一 (農研機構)
(SIG-SWO-038-06, 第 38 回, 2016 年 2 月 18 日)

選考理由:

本論文は, IT システムの活用による農業経営の効率化を目的とし, 農作業名称に関する標準語彙体系を提案するものである。構築されたオントロジーは表層的な用語分類にとどまらず, 記述論理を用いた明示的な意味定義に基づいており, その定義内容を用いた推論などの考慮が十分になされている。そのため, 総合運用性の高い農業情報の構築に向けた意義深い取り組みであるといえる。また, 農作業が持つ属性に着目し, それを細分化することにより階層構造を定義する過程は, 今後, オントロジー構築に取り組む研究者にとってモデルケースとなる有用なものである。以上のような観点から, 研究会優秀賞にふさわしい論文であるといえる。

7. SIG-AM インタラクティブ情報アクセスと可視化マイニング研究会

「情報アクセスにおける受動性と能動性: 音声対話によるニュース記事アクセス」

林 良彦 (早稲田大学), 藤江 真也 (千葉工業大学/早稲田大学), 福岡 維新, 高津 弘明, 小林 哲則 (早稲田大学)

(SIG-AM-11-05, 第 11 回, 2015 年 11 月 13 日)

選考理由:

本論文では, ニュース記事から, 利用者の反応の想定した発話計画を作成し, それに従って対話を行うことで, 利用者から見れば必要十分な情報アクセス, システムから見れば効率的な情報伝達が行える音声対話システムを提案している。このシステムの利用者は, 音声対話において自然と考えられる応答を返したり, 情報要求を提示したりすることで, 過不足ない情報アクセスを達成できる。本論文の貢献は, まずその優れた着眼にある。対話的な情報アクセスといった場合, 多くの研究が質問応答的な利用者主導に着目するが, ここでは, システム主導の情報提供をベースとして, そこに必要や関心に応じた利用者の主導権の行使を組み込むという新しい音声対話の形態が提案されている。そして, 実装においても, 構造に基づく文章要約技術, 音声対話に固有の発話に対する考慮, 計画立案など, 深い考察に基づく設計と実現がなされている。これらの点から本論文を今年度の研究会優秀賞にふさわしいとして選考した。

8. SIG-KBS 知識ベースシステム研究会

「アクティビティ評価情報と NMF の融合によるソーシャルメディアの信頼リンク予測」

松谷貫司（龍谷大学），熊野雅仁（龍谷大学），木村昌弘（龍谷大学），齊藤和巳（静岡県立大学），大原剛三（青山学院大学），元田浩（大阪大学）

（SIG-KBS-B502-02，第 106 回，2015 年 11 月 12 日）

選考理由：

本論文は、信頼リンクの生成およびアイテムのレビュー、他ユーザのレビューに対する評価の3種の機能を有するソーシャルメディアにおいて、直近数ヶ月間の行動から直後数ヶ月間における信頼リンクの生成を予測する問題を対象とし、ユーザの嗜好と信頼に対応した2つの潜在空間を利用する非負値行列分解(NMF)法を新たに提案すると共に、実験的にその性能を評価している。加えて、提案手法により得られる行列を利用し、嗜好と信頼の観点から、因子の解釈とユーザの行動分析を試みている。製品レビューと消費者報告のソーシャルメディアの一つであるEpinionsを対象とした評価実験では、信頼リンクとユーザの嗜好を利用したいくつかの既存手法と比較し、より高精度なリンク予測を達成しており、レビューに対する評価を適切に融合するという提案手法の妥当性及び有効性が示されている。また実験を通じ、Epinions データにおける信頼リンク生成の観点からの特徴的な性質を明らかにしている。評価はEpinionsのみを対象としており、実験数もやや少なく、また各種パラメタの設定などの課題も残されているが、リンク予測において利用可能な情報をNMFに取り込むための基盤的なアイデアと、得られる結果を利用した因子理解やユーザの行動分析への応用の可能性を示した点は高く評価でき受賞候補としてふさわしいと考える。

9. SIG-KSN 知識流通ネットワーク研究会

「セルフサービスシステムの実現に向けた回答データ作成手法」

竹内広宜，増田聡，宮本晃太郎（日本アイ・ビー・エム（株））

（SIG-KSN-018-06，第 18 回，2016 年 3 月 1 日）

推薦理由：

人工知能の産業界への応用が期待されるなか、顧客の問い合わせに自動対応する手法を提案する発表である。当該発表ではセルフサービスシステムの構築のために、質問と回答の対を集める手法を提案している。この手法は、顧客の行動段階と顧客が要求する知識をモデル化し、学習データを整理する点に特徴がある。当該発表では、銀行サービスにおける顧客からの質問と回答を対象にケーススタディを行い、人工知能による回答データ作成の効率化と品質向上についての可能性を主張している。顧客接点における人工知能の先駆的な応用しその可能性を明らかにした点は評価できる。よって研究会優秀賞にふさわしい発表であり、推薦する。

Can LOD Based Question Generation Support Work in a Learning Environment for History Learning?

Corentin JOUAULT¹, Kazuhisa SETA^{1,2}, and Yuki HAYASHI²

¹ Graduate School of Science, Osaka Prefecture University

² College of Sustainable System Sciences, Osaka Prefecture University

Abstract: The objective of this research is to build a system to support learners in self-directed learning of history in open learning space. The system supports learners with inquiry based learning using automatically generated questions. All the support provided by the system generated using the LOD. The evaluation of the system aims to verify the feasibility of LOD based support and its effectiveness. The results confirmed that LOD based support was feasible and that the question support improved engagement and historical deep thinking of the learners.

1 Introduction

When conducting self-directed learning in open learning space, one of the difficulties is that it requires learners to plan their learning in an unfamiliar domain. It requires learners to decide their learning objectives while they are learning. It is difficult for learners because they cannot easily judge the importance of the topics they encounter with their low knowledge of the domain. This problem is increased in open learning space because the quantity of information is much larger and the information is not organized with clear learning objectives. Learners can easily become overwhelmed and discouraged during learning because they cannot plan their learning. One way to eliminate the difficulty is prompting question generation and answering activities [Roth 96]. It contributes to lightening the difficulties of self-directing learning by lessening the planning activities that the learners need to perform. However, creating good inquiry questions requires an understanding of the domain and most learners cannot create good questions by themselves.

For this reason, we previously created a question generation function to create and adapt questions to learners. In a previous paper [Jouault 16], we demonstrated that it is possible to generate good quality history questions automatically in open learning space using Linked Open Data (LOD). A history professor judged that the quality of the automatically generated questions was on the same level as the ones generated by humans.

Another problem of self-directed learning is that learners must stay engaged in the learning task to have

fruitful learning outcomes. To raise the engagement of learners while keeping them motivated, the support should be adapted to the learners' interests and orient them to important information without forcing them.

Since the quality of the questions was demonstrated to be high enough to support learners, these questions need to be provided to learners in an adequate way to support their learning. Thus, we created a learning environment: the Semantic Open Learning System (SOLS) to support learners in self-directed learning of history with adapted question support.

The research question to be answered in this paper is whether automatic question generation support is feasible and can help learners in self-directed learning of history in open learning space. We will first describe the system and how the question generation function was used. Then, we will describe an evaluation of the effects of the system on learners. The evaluation aims to verify two points: (1) the usability of the system that uses question generation using the LOD and (2) the learning effects of the question generation support on basic knowledge acquisition, historical considerations and engagement in history learning.

2 SOLS: Semantic Open Learning System

2.1 System overview

Fig. 1 shows a screen image of the learning environment. It is composed of four windows:

- (a) **Document window:** It displays the learning material (Wikipedia document) the learner selected. All the

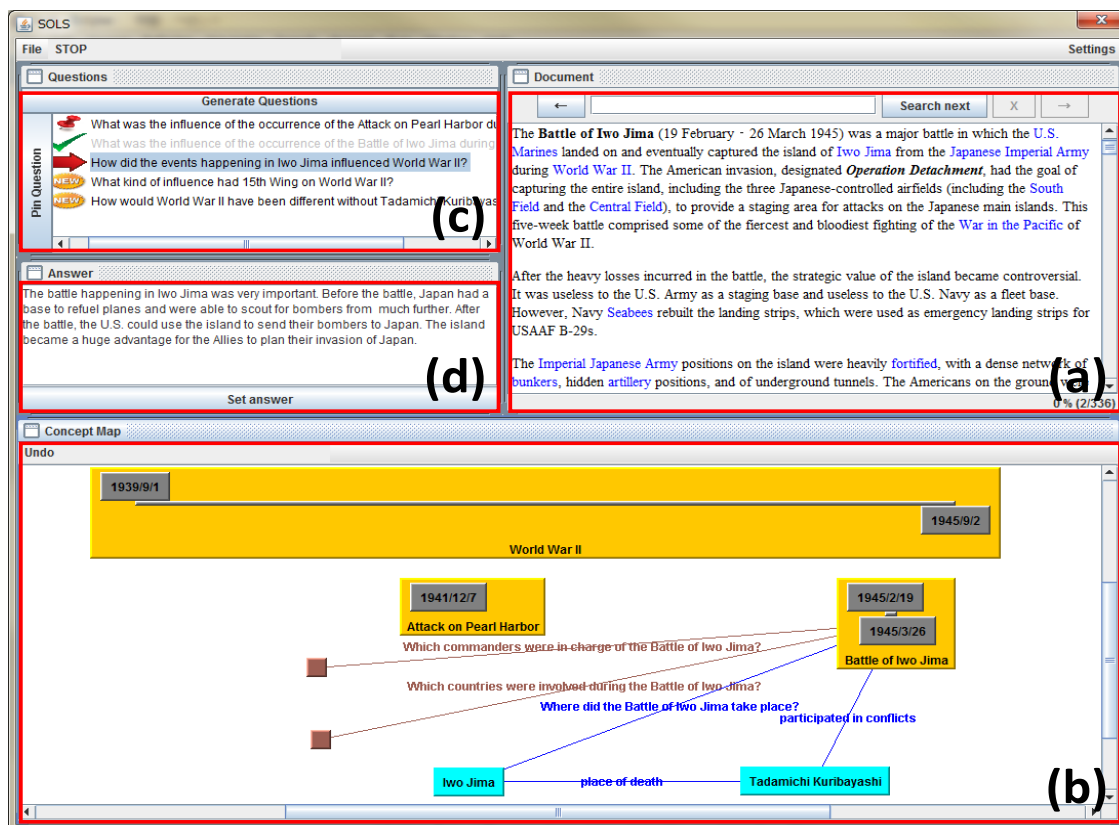


Fig. 1. System interface

links to a document about another concept appear in blue text. It provides usual Internet browser functions.

- (b) **Concept map window:** Learners use this window to manage the concept map. The concepts in the middle are events. Events are represented on an automatically generated timeline built using data from the LOD. Other concepts are colored in blue and can be moved freely by the learner. The lines between two concepts are relation with the type of relation written at the center of the line. The concept map window also displays “shallow” questions that are designed to support basic knowledge acquisition.
- (c) **Question window:** It displays a list of “deep” questions generated by the system designed to trigger deep historical thinking.
- (d) **Answer window:** Learners use this window to answer their active questions selected in (c). Learners can write their answer to the question in natural language.

When using SOLS, learners can request question to support their learning when they want. The questions appear either in the concept map (b) or the question window (c) depending on the learner’s request. By

choosing one of the generated questions, learners actually decide a learning objective to be reached. Learners then browse documents in window (a) to find an answer to the question. Finally, learners answer the chosen question in window (d) or in the concept map (b). By repeating this process, the learners build their understanding by performing inquiry based learning in open learning space.

The purpose of the concept map building task for learners is to make them to describe their understanding explicitly. Research showed that building a concept map deepens the understanding of learners [Nesbit 06]. Building a chronology also reinforces learners’ historical understanding [Stow 00].

In addition, the concept map is made fully machine understandable by using the LOD. Each concept added to the concept map has an ID that can be used to gather information on the LOD giving addition information about the concept to the system. Thus, it plays a key role to realize adaptive question generation that aims to deepen learners’ historical considerations and help learners set learning objectives. The machine understandable concept map makes it possible for the system to assess the learners’ knowledge and interests and adapt the generated questions accordingly.

2.2 Supporting learners with questions

All the questions are generated based on the LOD. The system can generate two sorts of questions depending on the learners' request: shallow questions in the concept map and deep questions in the question window.

First, on the learner's request, the system generates shallow questions appearing in the concept map to help the learners extend their basic knowledge. The generated questions will be displayed as a link between the concept and an empty node in the concept map. By displaying empty nodes, the system aims to trigger cognitive dissonances in the learner's mind. The empty node should make learners aware that they lack information and motivate them to fill the blank in their concept map. An example of concept map showing questions can be seen at the bottom of Fig. 1. These questions all require basic knowledge to be answered, such as a location for the questions "Where did the Battle of Iwo Jima take place?" The questions generated by this action are all questions about the concept selected by the learner. To answer the questions, the learners need to drag and drop the concept that answers it to the empty node and the system can verify the validity of the answer by using the data from the LOD.

Second, the question window (c) from Fig. 1 is designed to provide learners with adaptive advice generated based on the machine understandable concept map. Learners are provided with a list of deep questions in which they can select any that they consider interesting and try to answer it. In this situation, learners decide their learning objectives by choosing a question. The questions generated by the system were previously evaluated [Jouault 16] to be of a quality high enough to trigger historical thinking. The learners have access to good quality questions to direct their learning and their learning should be improved.

3 Experimental Setting

3.1 Objectives

The evaluation aims to verify four points:

1. The feasibility of LOD based support in a real learning scenario.
2. Whether the question support help learners develop their basic knowledge.
3. Whether the question support help learners develop their historical considerations.
4. Whether the question support can raise learners'

engagement in history learning.

Concerning the objective (1), we aim to verify whether the functions implemented in the system work smoothly to not disturb learning and are considered useful by learners. The amount of LOD stored in the system, to enable support in open learning space, is around 80 GB. Thus, we need to confirm that the system can work smoothly without frustrating the learners. The evaluation is the first evaluation involving the system and many of the functions implemented in the system that use the LOD have not been proved to be effective before. For this reason, the feasibility of the system should be verified to prove whether the system can successfully support learners using only automatically generated support created using the LOD.

Concerning the objective (2), the hypothesis to verify is that learners using the question support can develop a stronger basic knowledge than learners using the system without advice. We expect learners using the system to have more developed context knowledge because the shallow questions trigger a cognitive dissonance for the learners, they become aware of gaps in their knowledge and each question gives them direction to find the missing information.

Concerning the objective (3), the hypothesis to verify is that learners using the question support should perform better in task that require understanding and not only basic knowledge such as essay writing requiring historical considerations. We expect the deep questions to prompt learners to develop their historical considerations during learning.

Concerning objective (4), the hypothesis to verify is that the question support improves the learners' engagement and interest in history learning. We expect the engagement of learners to be raised because the questions give objectives to learners to facilitate their learning and it should minimize the difficulties of self-directed learning.

3.2 Procedure

Table 1 shows the timetable of the experiment. This evaluation involved 24 university students. They were separated in two homogeneous groups of 12 learners each to form the control and experimental groups depending on their results to a basic knowledge test about World War II (WWII). Before starting the evaluation, the learners are instructed to learn about World War I (WWI) on Wikipedia to make them aware of their history learning skills.

During the experiment, learners are firstly explained about good historical considerations in history learning

and inquiry learning. Then, learners are taught how to use the different functions of the system they will use with a demonstration and practice on the system for the WWI topic.

Before learning with the system, learners are instructed that they will have to write a report about their historical considerations on WWII after learning. Both groups are informed about the report and are instructed to study with that objective in mind.

The learning task is to learn about WWII including learning about one topic among two in detail. The topics to learn in details are two important events from World War II: the Attack on Pearl Harbor and the Battle of Iwo Jima. We set a self-directed learning situation where learners can choose freely according to their interests. The learners have 90 minutes (equivalent to a usual classroom lesson time) to perform the learning activities using the system.

During learning, both groups use the system to build their concept maps and browse documents. The difference between the groups is the availability of the question advice. The learners in the experimental group can request questions at any time to support their learning. The learners in the control group do not have access to

questions.

At the end of the learning phase, the learners have to:

1. Fill a questionnaire about the system to evaluate their feeling about the different functions of the system using the 5-grade Likert scale.
2. Answer a high-school level test about their chosen topic and WWII to evaluate their basic knowledge.
3. Write a report about their historical considerations to evaluate whether they performed deep historical thinking activities.

4 Results

4.1 Feasibility

To evaluate the results concerning the usability of the system, both groups should be considered since both groups used the system to build their machine understandable concept maps.

The category “System usability” of Table 2 shows the results of the questionnaire about the usefulness of the system. The learners from both groups gave good average scores on the Likert scale to the answer to the questions concerning the usability and usefulness of the system. The learners in both groups benefited from using the system while the average marks of the learners in the experimental group using the question advice were higher than the learners in the control group. It means that the system functioned as intended and was able help learners.

The category “Concept map usefulness” of Table 1 also shows that learners judged that building the concept map was useful. Most particularly, the learners felt that the timeline created using the LOD was particularly useful. It suggests that they felt the concept map and timeline building is helpful to improve their understanding of the topic.

The category “Questions’ quality” of Table 2 shows the results of the questionnaire for the questions only concerning the experimental group. It confirms the results of the evaluation of the questions which judged that the questions did not appear unnatural or nonsensical.

4.2 Effects on basic knowledge acquisition

The category “Shallow question usefulness” of Table 2 shows the results of the questionnaire concerning the shallow questions, which aim to improve learners’ basic knowledge. Learners judged that the shallow questions were useful and motivated learners in developing their basic knowledge and extend their concept map by

Table 1. Experiment timetable

Time	Con. (n=12)	Exp. (n=12)
Before experiment		
30 min	Study of WW I on standard browser	
10 min	Pre-Test about WWII for group formation	
Experiment		
5 min	Instructions: Good learning practices in history and inquiry learning	
35 min	<ul style="list-style-type: none"> • Demonstration of the system functions • Familiarization with the system 	
	System without question advice	System with question advice
90 min	Learning with the system (Subject: WWII + Attack on Pearl Harbor or Battle of Iwo Jima)	
	System without question advice	System with question advice
20 min	<ul style="list-style-type: none"> • Questionnaire • Basic knowledge test (Subject: WWII and chosen topic) 	
30 min	Report writing (Subject: Describe your historical considerations about WWII)	

Table 2. Questionnaire results

	Questionnaire item	Con.		Exp.	
		Avg.	SD	Avg.	SD
System usability	Do you think using the system helped you understand the topic?	3.92	0.67	4.25	0.62
	Do you feel that the system made it easier to learn?	3.50	1.09	3.92	0.79
Concept map usefulness	Do you think the timeline part of the concept map helped you understand the events?	4.25	0.75	4.50	0.67
	Do you think building a timeline including the context is a good way of learning history?	4.33	0.65	4.50	0.67
Interest in history	Are you interested in learning history?	3.00	1.28	3.58	0.90
	Do you want to learn more about history?	3.08	1.24	3.50	1.00
	Do you think using the system made you want to learn more?	3.08	1.24	3.67	0.65
	Do you feel using the system made you more interested in learning about history?	3.17	1.19	4.00	0.74
Interest in learning method	Do you think it is easier to perform inquiry based learning using the system?	3.33	0.78	4.25	0.62
	Do you think being aware of the questions is helpful to learn history?	3.67	0.78	4.58	0.51
Questions' quality	Do you think the questions seemed unnatural?	-	-	2.67	1.44
	Do you think the questions were nonsensical?	-	-	2.42	1.51
Shallow questions usefulness	Do you think the questions in the concept map made you realize what knowledge you needed to develop?	-	-	4.08	0.67
	Do you think the empty node that appears which each question in the concept map made you want to answer them?	-	-	4.00	0.74
	Do you think the questions not answered in free-text were useful?	-	-	3.75	0.75
Deep questions usefulness	Do you think the questions proposed by the system were about topic you were focusing on at the time?	-	-	3.42	0.79
	Do you think the questions answered in free-text were useful?	-	-	4.33	0.65

triggering cognitive dissonances. This confirms that the shallow question worked as intended.

However, even though the learners judged that the shallow questions were useful, the results of both the test about general knowledge of WWII and the topic chosen both show no significant difference between both groups for the tests about the chosen topic and about WWII. The question support in the short term use of the study does not seem to have a significant effect on the learning outcomes of the learners.

4.3 Effects on historical considerations

The category “Deep questions usefulness” of Table 2 shows the results of the questionnaire for the questions concerning the deep questions. The deep questions were judged useful by the learners. The learners also felt that the questions were adapted to their interests and knowledge as the questions were related to what they were studying. It means that the system was able to generate adapted deep questions to help the learners.

The reports were graded by a history professor with scores from 1 to 5. The results of the report show that there is little difference by the short term use between the grades of the reports from both groups (Control: 2.08,

Experimental: 2.33).

More detailed analysis of the content performed by the history professor that graded the reports is shown in Table 3. The evaluator categorized the contents of the reports in 5 categories:

1. Personal feeling: the report describes the learners’ personal feeling about the events.
2. Fact enumeration: the report is mostly a list of facts described without reflection by the learner.
3. Lesson learned: the report describes the lesson that should be learned from the events and makes the connection between the events and the current situation.
4. Historical considerations: the report describes the learner’s historical considerations about the topic. The report contains the results of deep historical thinking from the learners.
5. Irrelevant: the report’s contents cannot be categorized in another category or are confusing.

Some reports had contents that could be categorized in two different categories (3 for control group, 2 for experimental group). These reports are included in the count for both categories of the Table 3 in these cases. From 1 to 4 (excluding irrelevant content in 5), the history

Table 3. Categorization of reports content

	Personal feeling	Fact only	Lesson learned	Historical considerations	Irrelevant
Control (3 reports in 2 categories)	2	8	3	1	1
Experimental (2 reports in 2 categories)	0	3	2	6	3
Average number of questions answered (Experimental group only)	-	3.33	5.00	5.17	2.33

professor judged that the quality became higher from the viewpoint of historical considerations.

The detailed results show that even though the grades were similar, the content categories were different. Most learners in the control group (not using the question advice) wrote reports that are mostly enumerations of facts with little historical considerations. On the other side, many of the learners in the control group wrote reports containing deep historical considerations. The results also show that the more deep questions were answered by learners, the more their report is categorized in higher quality. It suggests that the deep questions prompt their historical considerations as intended.

4.4 Effects on engagement

The category “Interest in history” of Table 2 shows that learners in the experimental group showed more interest in history learning after using the system. The results show that the question generation support has a potential to raise engagement of the learners in history learning as the average mark for learners in the experimental group is higher. Since the learners in self-directed learning can choose to stop learning at any time, it is important to keep learners engaged in the learning task. Engaged learners will spend more time for learning and will develop their knowledge further.

The category “Interest in learning method” of Table 2 shows that using the system made the learners more aware of the usefulness of questions in self-directed learning of history. The results that the learners felt that the questions provided by the system helped them be aware of the importance of questions when conducting self-directed learning of history.

5 Concluding Remarks

The results of the evaluation showed that supporting learners in open learning space using support automatically generated using the LOD is feasible and can be useful for learners. Most learners judged that using the

system was useful and that it helped them learn about history.

Even though the short term use of the question generation support did not have a significant effect on the learners’ knowledge acquisition, it had an effect on the development of historical considerations of learners and motivated them to perform deep historical thinking to develop their opinion.

The question support also had an effect on the engagement of learners. Engagement has strong effects in self-directed learning because being engaged and interested in the topic lead learners to study for a longer time and learn about more topics

References

- [1] Hmelo-Silver, C. E., Duncan, R. G. and Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, Vol.42, No.2, pp. 99-107.
- [2] Jouault, C., Seta, K. and Hayashi, Y. (2016, to appear). Content-Dependent Question Generation using LOD for History Learning in Open Learning Space. *New Generation Computing*. Springer International Publishing.
- [3] Nesbit, J. C. and Adesope, O. O. (2006) Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, Vol. 76, No.3, pp. 413-448.
- [4] Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching*, Vol.33, No.7, pp. 709-736.
- [5] Slotta, J. D. (2004). The web-based inquiry science environment (WISE): Scaffolding knowledge integration in the science classroom. In *Internet Environments for Science Education*, pp. 203-232.
- [6] Stow, W. and Haydn, T. (2000). 7 Issues in the teaching of chronology. In *Issues in History Teaching*, p. 83. Routledge.
- [7] Zimmerman, B. J. and Schunk, D. H. (Eds.). (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*. Routledge.

ニューラルネットワーク物語生成モデル

A Neural Narrative Model

浅川 伸一

Shin Asakawa

東京女子大学情報処理センター

Center for Information Science, Tokyo Woman's Christian University

The author tried to apply recurrent neural network models to generating narratives. Inspired by recent advances of recurrent neural networks and successors, sequences, or sequences of sequences, might be generated stories in terms of recurrent neural networks. However, the roles of attention and primed stimulus remained to be unsolved.

1. はじめに

ディープラーニング (深層学習) の進展は、認識、判別のみでなく、制御、生成を視野に入れた応用がある。2013 年以降を考えても、ディープドリーム (<http://deepdreamgenerator.com/>)、画像を芸術家風に変形させるディープアート [6]、画像にキャプションを付ける Neural Image Captioning (NIC, hereafter)[22]、手書き文字生成 [9]。ことば周辺領域に限定しても従来の統計的機械翻訳を BLEU で凌駕した [21]。数式処理 [24]、プログラムコード生成 [25]、アドベンチャーゲームなどもある [23]。ニューラルチューリングマシンは並べ替えアルゴリズムを習得した [10]。これらは系列生成の変形であり、従来手法の再利用、再評価でもある。本稿では上述の論文で用いられるようになった技術を生成の側面から整理を試みた。

2. リカレントニューラルネットワーク

リカレントニューラルネットワーク (Recurrent Neural Networks) とは連続情報を扱うニューラルネットワークである。音声認識、機械翻訳、動画像認識など予測次元数が可変な場合に用いられる。時刻 t の隠れ層の状態が時刻 $t+1$ の隠れ層への入力となる。

Mikolov ショック (2010 年)[18] 以来リカレントニューラルネットワークが言語情報処理で再評価された。

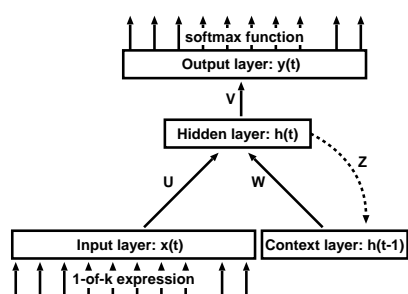


図 1: Mikolov の単純再帰型ニューラルネットワーク [18] を改変

Mikolov のアーキテクチャは単純再帰型ニューラルネットワークである [4, 5]。入力系列を x 、出力系列を y とすれば単

連絡先: 167-8585 杉並区善福寺 2-6-1, asakawa@ieee.org

純再帰型ニューラルネットワークは

$$y(t) = f(Ux(t) + Wh(t-1)) \quad (1)$$

と表現される。ある時刻における出力は入力と一時刻前の状態に依存する。しかし、一時刻前の状態は、その時の入力とさらに一時刻前の状態とに依存するので単純再帰型ニューラルネットワークであっても長期的な過去の依存関係、系列情報、因果関係を表現可能である。系列長が可変なデータに対しても適用可能な点に特徴がある。

学習については一時刻前の状態を文脈層にコピーするので、中間層から見ると入力情報が増えた以外に、通用のフィードフォワード型のニューラルネットワークの学習と相違はない。任意の目標関数を各結合係数で微分したバックプロパゲーションによる学習が行われてきた。

一般にリカレントニューラルネットワークを考えた場合、学習については系列情報をどこまで考慮するかにより BPTT と RTRL の 2 つのモデルが提案されてきた。BPTT と RTRL とを統一して表記することも可能である [11]。

Mikolov は長「短期記憶」を使っていなかった [18]。しかし、近年のモデルでは長「短期記憶」によるリカレントニューラルネットワークの性能向上と収束時間も貢献している (2.1 節)。

ただしリカレントニューラルネットワークの多層化については評価が困難である。ここでは以下の 3 点を検討した。

1. 長「短期記憶」
2. 注意
3. プライム

これらの意味合い、及び、評価は定まっているとは言えない。例えば、画像認識、音声認識で state of the art である畳み込みニューラルネットワークで用いられる (三種の神器) :

1. Max pooling
2. Dropout
3. Rectified Linear Units (ReLU 整流線型ユニット)

ほど明確ではない。ただし三種の神器だけを信じる理由もない。例えば、最大値だけプールして他の情報を廃棄する以外にも方法はある [16]。Dropout は単なる正則化ではない。整流線型ユニット以上に性能が向上する方法も提案された [12]。

2.1 長「短期記憶」

長「短期記憶」(Long-Short Term Memory)[14, 7] はリカレントニューラルネットワークであるが、直近の情報だけに依存するのではなく、遠い過去、すなわち長期間に渡って記憶を活用することが可能である。これは直近の記憶をゲートによって開閉することで可能となる (図 2)。

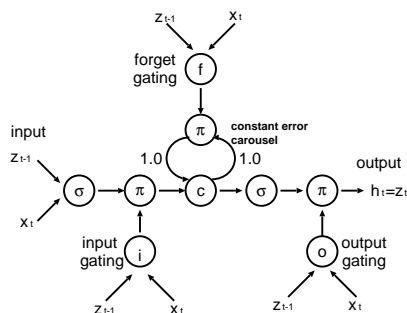


図 2: 長「短期記憶」

図 2 では情報は左から右に流れる。図 2 が一つのユニットを構成する基本単位である。この基本単位を構成要素としてニューラルネットワークが構成される。図 2 中央の c が記憶ユニットである。 c の前後二つのシグマバイユニットすなわちゲートが配置されそれぞれ入力、出力ゲートと呼ばれる。 c の自己フィードバックのゲートは忘却ゲートと呼ばれる。 c の出力すなわち過去の状態からの再帰結合は恒等関数 $f(x) = x, \forall x$ であるため 1.0 と表記した (CEC: constant error carousel)。 c への入力 は CEC と忘却ゲートの出力との積である。

出力ゲートと入力ゲートの役割は、入出力量をスケールリングすることで CEC にどの程度の誤差を学習するかを指示するように振る舞う。入力ゲートによるスケールリングと出力ゲートのスケールリングにより誤差を開放するか否かの情報を記憶ユニットが学習することになる。忘却ゲートは CEC に対してすなわち一時刻前の自身の保持している情報を利用するか否かを定めるためのゲート開閉を行う。

時刻 t における入力: i_t , 忘却ゲート: f_t , 出力ゲート: o_t , 長「短期記憶」ユニットの出力: h_t , 記憶セル: c_t とすれば長「短期記憶」のユニットは以下のように表わされる。

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i), \quad (2)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f), \quad (3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o), \quad (4)$$

$$g_t = \phi(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t, \quad (6)$$

$$h_t = o_t \odot \phi(c_t) \quad (7)$$

W は関連する結合係数行列, σ はロジスティック関数, ϕ は \tanh である。 \odot はシグマバイユニットすなわち積を表している。ゲートの出力に \tanh を用いているため、正確にはゲートの開閉 $[0, 1]$ ではなくゲートの反転 $[-1, 1]$ である。

リカレントニューラルネットワークを時間展開し, $t-2$ の過去からの情報を利用する様子を図 3 に示した。

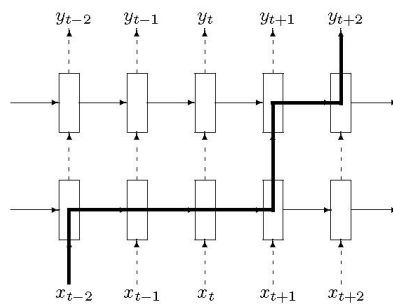


図 3: 長「短期記憶」の説明 [26] より

2.1.1 勾配消失問題, 勾配爆発問題

バックプロパゲーション法においては任意の課題における特定の出力層ユニットの誤差が下位層の全ユニットに伝播する。従って多層化されたニューラルネットワークで, かつ, 活性化関数をシグモイド関数 ($\sigma(x) = (1 + \exp -x)^{-1}$) とした場合, 誤差関数を各結合係数で微分した値にシグモイド関数の微分が入る (式 (8))

$$\frac{d\sigma(x)}{dx} = \sigma(x)(1 - \sigma(x)). \quad (8)$$

$0 \leq \sigma(x) \leq 1$ であるので $d\sigma(x)/dx$ は微分するたび, すなわち層を下るたびに小さくなる。勾配消失問題の一因はここにもある。勾配消失問題を明確に記述したのは Bengio[2] であるが, 以前から問題であった [13]。

一方, リカレントニューラルネットワークは系列が長くなると再帰が深くなる。伝播した勾配が発散する。最小化すべき損失関数 L をパラメータ θ で時間窓 t_w までの勾配を計算すると

$$\frac{\partial L}{\partial \theta} = \sum_{t \in t_w} \frac{\partial L}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial h_t} \frac{\partial h_t}{\partial \theta^{(t-1)}}, \quad (9)$$

ここで $\partial h_{t-1}/\partial \theta^{(t-1)}$ は中間層 h での時刻 t におけるパラメータ θ の微分である。階層が深くなると $\partial h_t/\partial h_{t-1}$ のヤコビアン最大の特異値が 1 より小さければ勾配消失問題となり, 1 より大きければ勾配爆発問題となる [2]。リカレントニューラルネットワークの学習においては式 (9) で再帰的に勾配を計算した場合, ヤコビアンの最大特異値に応じて指数関数様に勾配が変容することになる。

2.2 注意

前述の長「短期記憶」は直近の入力情報をゲートの開閉によって制御することで系列情報処理を行う方法である。一方, 直近の出力情報から記憶状態を制御する方法として注意が導入される場合がある。

Graves はリカレントニューラルネットワークを用いて予測と制御を行った [9]。この際, 中間層ユニットの出力にガウシアン関数で時間窓を設けた。これにより中間層ユニットの状態は出力に依存して直後の状態を制御できるようになる (図 4)。

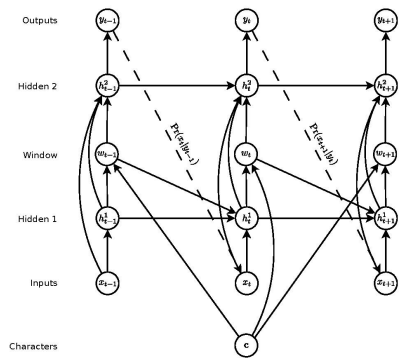


図 4: Graves の注意 [9] より

一方, Mnih らも出力情報を次刻の中間層ユニットの制御に利用した (図 5 の点線)。

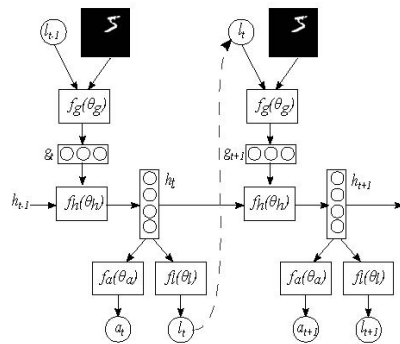


図 5: Mnih の注意 [19] より

さらに Bahdanau らは出力側の S から注意と名付けたデコーダ側の情報をエンコーダ側の制御に用いた (図 6)。

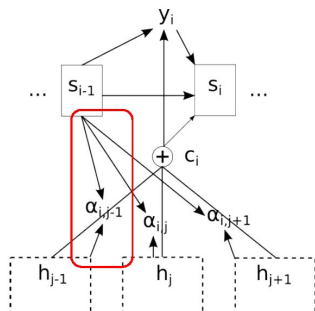


図 6: Bahdanau の注意 [1] より

Bahdanau らの注意は softmax 関数である (式 (10))。

$$\alpha_{i,j} = \frac{\exp(\alpha_{i,j})}{\sum_k \exp(\alpha_{i,k})} \quad (10)$$

である。

2.3 プライム

系列生成のためには、種となる情報が必要となる。Gaves はこれをプライムと呼んだ。ディープドリームやディープアートにおいては元画像である。

グレース (A. Graves) の手書き文字生成 [9] モデルは再帰結合を有する混合ガウシアン関数である。従って、ある系列を予測する場合、核関数であるガウシアン関数に対して平均 μ にバイアス b を加える事で系列を制御した。後述の A4 節を参照のこと。

一方 Gatys らによる画風の変換においては原画像と特徴ベクトルとの差異を勾配降下法で解くことにより画風を変換させた (A6 節)。

3. 実験

以上の 3 点, 長「短期記憶」, 注意, プライムを考慮して, 青空文庫を材料に文章生成を行った。

3.1 結果

プライムを与えず, 初期値を語彙から選んで生成させた場合を示した。

坂口安吾, 墮落論:

変り易いから「節婦は二夫に見えず、たまさか気づく者があって、私の一生を正視するに堪えないような気がしておらず、たまさか気づく者があったに相違ない。けれども、外の事柄に就ても人間自体をどう為しうるものであり、ただ表面的な真理や自然法則だけであるか。之も亦歴史の前ではないようだ。実際、泥棒すらも持ちうるのかも知れぬと思っていた。偉大な破壊が好きであったが、二十年の四月四日という日、私の姪の場合であった。無心であり、ただそれを防ぐことによって永遠の隆盛（それは人間の弱点を抑える必要があっても、考えることが必要で、個々の政治を貫く巨大な独創をもち、意慾をもち、やむべからざる歩調をもって大海の波の如くに歩いて

北原白秋, お月様いくつ:

まあいとお月さまの中には、いつも兎が杵《きね》をもつて餅を搗《つ》いてといふのがありませう。ほんとに金屏風でなくては、あの若い小母さまには似合はないでせうね。いかにも昔のお江戸の子供が謡つたやうでせう。氣象《きしやう》これは、空のけしきが其のままに歌はれてゐます。小さい和尚さんは白い星か薄《うす》い霧のやうな星の雲かでせう。滑橋《すべりばし》もさうした雲のながれでせう。

狸の子はさっきのようにとんとん叩きながら時々頭をまげてセロに耳をつけるようになりました。そしておしまいまで来たときは今夜もまた東「ああ夜が明けたぞ。どうもありがとう。」狸の子はさっきのようにとんとん叩きながら時々頭をまげてセロに耳をつけるように

しました。「ああ夜が明けたぞ。どうもありがとう。」狸の子はさっきのようにとんとん叩きながら時々頭をまげてセロに耳をつけるようにしました。「ああ夜が明けたぞ。どうもありがとう。」狸の子はさっきのようにとんとん叩きながら時々頭をまげてセロに耳をつけるようにしました。すると猫は肩(かた)をまるくして眼をすぼめてはいましたが口のあたりでにやにやわらって云いました。はじめました。

4. 考察

正当にモデルを評価するためには長「短期記憶」であるべきなのか、多層リカレントニューラルネットワークにするなら何層が良いのか、注意の有無、プライムの与え方を比較検討すべきであった。今回は例示に留まった。またリカレントニューラルネットワークに対して dropout や正則化を行うべきか否かについても議論が分かれる (Zaremba[26] を参照)。

5. 結語

本稿では物語生成の可能性を示した。考慮すべき点として3点、長「短期記憶」、注意、プライムを指摘した。これらの詳細な記述と関係は今後の課題であるが可能性を示したと考える。

A 関連研究

A1 ニューラル画像脚注付け

Vinyals らは画像から説明文を生成 [22] した。静止画像から説明文を生成する。画像認識には畳み込みニューラルネットワーク文章生成にはリカレントニューラルネットワークを用いている。

A1.1 グーグルの方法

視覚畳み込みニューラルネットワークから言語生成リカレントニューラルネットワーク(長「短期記憶」)へと接続するモデルである [22]。

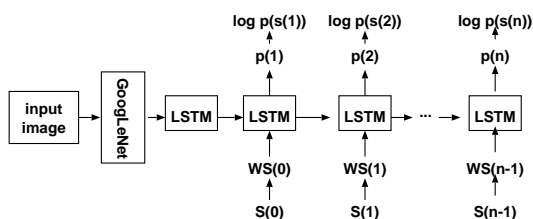


図 7: 左から GoogLeNet の出力が長「短期記憶」への入力となる。全ての長「短期記憶」は係数行列を共有する。[22] より

画像 I から説明文 S を生成する。すなわち推定は $S^* = \operatorname{argmax}_S p(S|I)$, ($S = \{S_0, S_1, \dots\}$) による。ここで

$$p(S|\theta) = \prod_{t=0}^N p(S_t|S_0, S_1, \dots, S_{t-1}; \theta) \quad (11)$$

である。

$$x_{t-1} = \operatorname{CNN}(I) \quad (12)$$

$$x_t = W_e S_t, \quad t \in \{0, \dots, N-1\} \quad (13)$$

$$p_{t+1} = \operatorname{LSTM}(x_t), \quad t \in \{0, \dots, N-1\} \quad (14)$$

学習は文 S の対数尤度の最大化:

$$L(I, S) = - \sum_{t=1}^N \log p_t(S_t). \quad (15)$$

による。

A1.2 スタンフォード大の方法

畳み込みニューラルネットワーク+リカレントニューラルネットワーク [15] は領域切り出し(対象の領域割り当て)は画像 k と文 s について

$$S_{kl} = \sum_{t \in g_l} \sum_{i \in g_k} \max(0, v_i^T s_t), \quad (16)$$

を考える。 g_k は領域切り出しをした各画像断片であり, g_l は文 S の断片(文節)である。または上式を単純化した

$$S_{kl} = \sum_{t \in g_l} \max_{i \in g_k} v_i^T s_t, \quad (17)$$

でも良い。

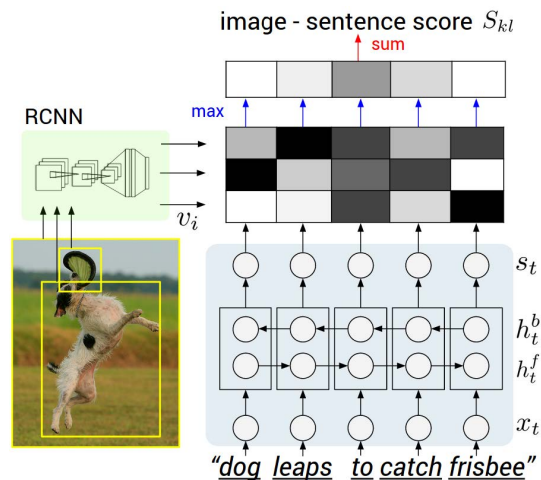


図 8: 2つの対象「犬」と「frisbee」。[15] より

図 8 左は画像例。画像断片が v_i で文 s_j との対応関係を $y_{ij} = \operatorname{sign}(v_i^T s_j)$ として扱う。

リカレントニューラルネットワークは一つの単語を入力とし、それ以前の時刻の状態を文脈として扱う。画像断片と文パーツとの関連を潜在変数とするマルコフ確率場 MRF を評価する。

$$E(a) = \sum_{j \in N} \phi_j^U(a_j) + \sum_{j \in N-1} \phi_j^B(a_j, a_{j+1}) \quad (18)$$

$$\phi_j^U(a_j = t) = v_i^T s_t \quad (19)$$

$$\phi_j^B(a_j, a_{j+1}) = \beta \mathbf{1}[a_j = a_{j+1}] \quad (20)$$

A1.3 UCB の方法

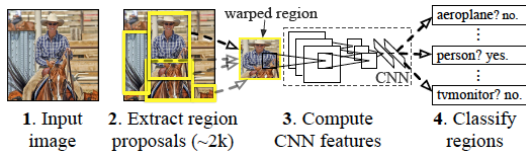


図 9: Girshick et al. R-CNN[8] 図 1 より

Girshick ら [8] はボトムアップに小領域の特徴を抽出して畳み込みニューラルネットワークへの入力とし、サポートベクターマシン [3] によって領域の分類を行った。

A2 ニューラルチューリングマシン

記号操作を抽象化して考えれば、言語情報処理とコンピュータプログラムと人間の作業記憶とは強く関連する。Graves らは Neural Turing machines:ニューラルチューリングマシンを提え、その枠組を示した [10]。ニューラルチューリングマシンは「アルゴリズム」を扱うことが可能である

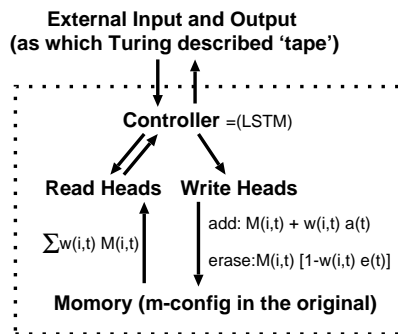


図 10: ニューラルチューリングマシンの概念図

ニューラルチューリングマシンの要点も系列制御であり、コントローラに長「短期記憶」を用いている。

A3 S2S

現時点で最高性能 (SOTA) モデルである Sutskever ら [21] のモデルを S2S と表記する。S2S は長「短期記憶」層を 4 層にしたモデルである。加えて、ソース言語の語順を逆転させている。ターゲット言語は逆順にしない。すなわち 'a', 'b', 'c' というソース言語を ' α ', ' β ', ' γ ' に翻訳する際、入力系列として (c, b, a) を入力とし、(α , β , γ) を出力させると成績が向上した。ソース言語を逆順にして入力すると性能が向上した理由に浮いては長「短期記憶」が処理しやすいからと書いてある。Sutskever らの方法を図 11 に示した。(http://www.cs.toronto.edu/~ilya/rnn.html)

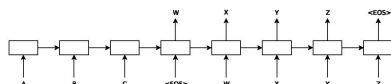


図 11: Sutskever らのリカレントニューラルネットワーク言語モデル [21] より

A4 手書き文字生成

Graves による手書き文字のデモ (<http://www.cs.toronto.edu/~graves/handwriting.cgi>) に従って Kotoba Engineering at Morioka の生成例を図 12 に示した。

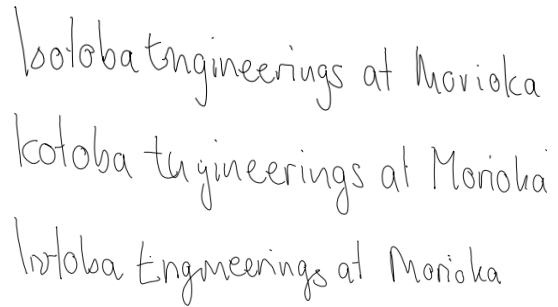


図 12: Kotoba Engineerings at Morioka

A5 ディープドリーム

[17] (<http://deepdreamgenerator.com/>) を参考にした。後述するディープアートと同様に学習済みのネットワークに原画像由来の情報を加工したと思われる。



図 13: マチュピチュの夢



図 14: 天安門前広場の夢



図 15: 真実の口の夢

A6 ディープアート

Gatys らによる画風の変換では VGG[20]*1 が用いられた [6]。Gram 行列 $G^l \in \mathcal{R}^{N_l \times N_l}$ すなわち G_{ij}^l は第 l 層 の特徴地図ベクトル i と j との内積が用いられた (式 (21))。

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l. \quad (21)$$

所与の画像のテクスチャを生成するために、白色雑音から勾配降下法を用いた。Gram 行列と原画像との平均自乗距離を最小化するように a を原画像, x を Gram 行列とする。 A^l と G^l とは第 l 層の表象だとする。この層の誤差は

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2, \quad (22)$$

と定義される。これを用いて全体の誤差は、

$$L_{style}(a, x) = \sum_{l=0}^L w_l E_l, \quad (23)$$

各層毎の貢献を重み係数 w_l と表記した。各層毎の誤差 E_l を活性化値で微分して、

$$\frac{\partial E_l}{\partial F_{ij}^l} = \begin{cases} \frac{1}{N_l^2 M_l^2} \left((F^l)^T (G^l - A^l) \right)_{ji} & \text{if } F_{ij}^l > 0 \\ 0 & \text{if } F_{ij}^l < 0 \end{cases} \quad (24)$$

を得る。これを用いて最終的な損失関数：

$$L_{total}(p, a, x) = \alpha L_{content}(p, x) + \beta L_{style}(a, x), \quad (25)$$

を得る。これを最小化することで画風を変換させた。図 16 に畳み込みニューラルネットワークの各層毎の変換の様子を示した。

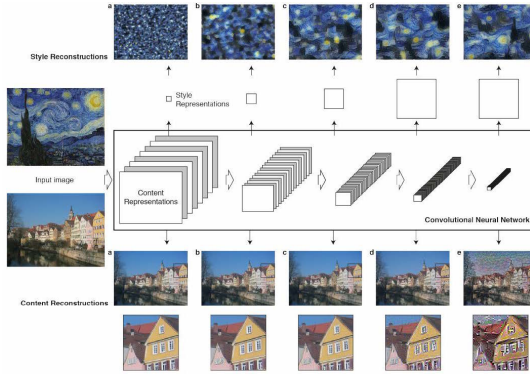


図 16: ディープアートのカラクリ [6] より

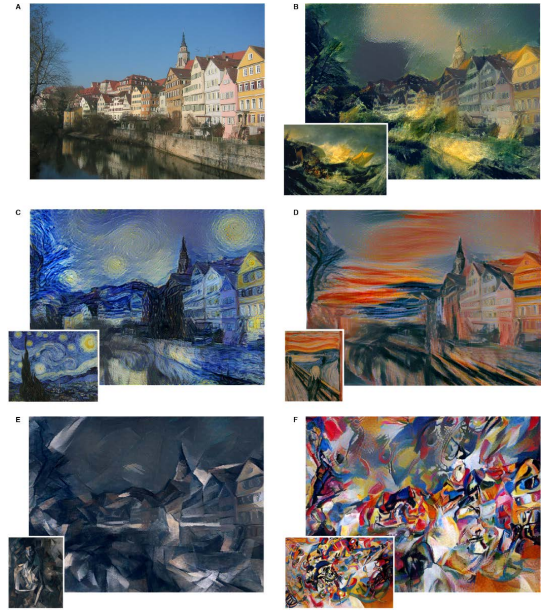


図 17: ディープアートの生成画像 [6] より



図 18: カジンスキー画風変換例 [6] より

参考文献

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In Yoshua Bengio and Yann LeCun, editors, *Proceedings in the International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.
- [2] Yoshua Bengio, Nicolas Boulanger-Lewandowski, and Razvan Pascanu. Advances in optimizing recurrent networks. *arXiv*, 2012.
- [3] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir Naumovich Vapnik. A training algorithm for

*1 VGG は ILSVRC2014(大規模画像認識チャレンジ)に出場したチーム名に由来する。現在は DeepMind

- optimal margin classifiers. In D. Haussler, editor, *the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM press, 1992.
- [4] Jeffrey L. Elman. Finding structure in time. *Cognitive Science*, 14:179–211, 1990.
- [5] Jeffrey L. Elman. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7:195–225, 1991.
- [6] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style. *arXiv*, 2015.
- [7] Felix A. Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with LSTM. *Neural Computation*, pages 2451–2471, 2000.
- [8] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of Computer Vision and Pattern Recognition Conference (CVPR)*, Columbus, Ohio, USA, 2014.
- [9] Alex Graves. Generating sequences with recurrent neural networks. *arXiv*, 2013.
- [10] Alex Graves, Greg Wayne, and Ivo Danihelka. Neural Turing machines. *arXiv/cs*, 2014.
- [11] Thomas Hanselmann, Anthony Zaknich, and Yianni Attikiouzel. Connection between BPTT and RTRL. *Computational Intelligence and Applications*, 1:97–102, 1999.
- [12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. Technical report, Microsoft, 2015.
- [13] Sepp Hochreiter, Yoshua Bengio, Paolo Frasconi, and Jürgen Schmidhuber. Gradient flow in recurrent nets the difficulty of learning long-term dependencies. In S. C. Kremer and J. F. Kolen, editors, *A Field Guide to Dynamical Recurrent Neural Networks*. IEEE press, 2001.
- [14] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9:1735–1780, 1997.
- [15] Andrej Karpathy and Li Fei-Fei. Deep visual-semantic alignments for generating image descriptions. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, June 2015.
- [16] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *arXiv:1312.4400v3*, 2014.
- [17] Aravindh Mahendran and Andrea Vedaldi. Understanding deep image representations by inverting them. 2015.
- [18] Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan “Honza” Černocký, and Sanjeev Khudanpur. Recurrent neural network based language model. In Takao Kobayashi, Keiichi Hirose, and Satoshi Nakamura, editors, *Proceedings of INTERSPEECH2010*, pages 1045–1048, Makuhari, JAPAN, September 2010.
- [19] Volodymyr Mnih, Nicolas Heess, Alex Graves, and Koray Kavukcuoglu. Recurrent models of visual attention. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2204–2212. Curran Associates, Inc., 2014.
- [20] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun, editors, *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.
- [21] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems (NIPS)*, pages 3104–3112, Montreal, BC, Canada, 2014.
- [22] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. In *Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015.
- [23] Jason Weston, Sumit Chopra, and Antoine Bordes. Memory networks. In Yoshua Bengio and Yann LeCun, editors, *Proceedings in the International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.
- [24] Wojciech Zaremba, Karol Kurach, and Rob Fergus. Learning to discover efficient mathematical identities. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, pages 1278–1286, Montreal, BC, Canada, 2014. Curran Associates, Inc.
- [25] Wojciech Zaremba and Ilya Sutskever. Learning to execute. In Yoshua Bengio and Yann LeCun, editors, *Proceedings of the International Conference on Learning Representations, (ICLR)*, San Diego, CA, USA, 2015.
- [26] Wojciech Zaremba, Ilya Sutskever, and Oriol Vinyals. Recurrent neural network regularization. In Yoshua Bengio and Yann LeCun, editors, *Proceedings of the International Conference on Learning Representations, (ICLR)*, San Diego, CA, USA, 2015.

株式掲示板におけるユーザ行動異常検知を用いた 相場操縦発見手法に関する研究

宮崎 邦洋^{1*} 松尾 豊¹

¹ 東京大学工学系研究科技術経営戦略学専攻

1 背景

現在，日本市場において個人投資家の存在が大きくなっている¹．この流れは日本の経済にとって好ましいものであり，政府としては個人の資金をさらに市場に向わせたい旨をその提言の中で述べている²．

1.1 Web 上における風説の流布

一方，インターネットの普及により，ソーシャルメディア上で風説の流布が行い易くなったため，個人投資家が安全に投資を行えなくなるような要因も増していると言える．実際掲示板の風説の流布によって告発されたケースも複数ある³．

そのような行為をする者を相場操縦士や仕手師と呼ばれるが，この犯罪行為の標的にされやすいのもまた個人投資家である．何故なら個人投資家は機関投資家と比較して情報劣位にあり，非合理的な投資行動を取りやすいからである [1, 2]．しかし先述したように，東京市場の活性化という観点においては個人投資家の存在はとても大きいものである．現在の個人投資家売買比率増加の傾向を持続させるためにも，市場の公正性・信頼性を保つことはより重要になってくる．

もちろん証券取引等監視委員会は，市場の監視を行っているが，大量の情報が生成されるソーシャルメディアにおいて，すべての書き込みを監視することは容易ではない．

そのような問題に対して，Web マイニングの技術が有効である．Web マイニングではソーシャルメディア上に所謂ビッグデータを抽出・構造化・分析することで，現実世界のデータからでは得難い知見をより簡単に得ることができる．

この技術を用いることで，かつて困難であったソー

シャルメディアの監視を効率的に行えるようになる．

以上の内容を踏まえ，本研究では，インターネット株式掲示板において相場操縦を発見することを目的とする．そのため，掲示板におけるユーザの行動を分析し，相場操縦行為を行った可能性のあるユーザを発見する手法を提案する．

本研究が提案する手法は，急騰急落直前のユーザの行動が，通常時と比べどれくらい異常であるかを測定し，その異常具合をスコア化・ランク付けするものである．提案手法の使用用途については，金融庁など相場操縦を発見する主体を支援するシステムなどが想定される．

2 関連研究

2.1 相場操縦の研究

相場操縦とは一般に裁定取引機会の創出と利用と定義されている [3]．この相場操縦に関しては，主に直接的に取引戦略によって行われるもの（見せ玉等）と，間接的に情報操作などによって引き起こされるものがある．

本研究は後者に関係するものだが，[4] では，1600 年から今日の情報化社会に至るまでにおいて，技術の発展によって株式銘柄に関しての情報が世の中に容易に拡散されやすくなり，それによって投資家が影響を与えられやすくなっていることが示されている．その技術の中でもインターネットの発達により，投資家が世界中に情報を発信できることや，世界中から情報を得ることができるようになったことに大きく貢献している．

インターネットにおける風説の流布において株式掲示板はその主たる媒体であるが，その株式掲示板を使用した相場操縦の研究については，Delort らがオーストラリアの株式掲示板に関する研究を行っている [5]．この研究では，HotCopper⁴ という掲示板について，2008 年のオーストラリア証券取引所上場銘柄 1825 社に関して，100 万余りの書き込みの分析から，相場操縦を検証し

*E-mail : miyazaki@weblab.t.u-tokyo.ac.jp

¹NTT データ経営研究所による調査
<http://www.jpex.co.jp/files/tse/news/31/b7gje6000003lhav-att/b7gje6000003lhjw.pdf>

²財務省「金融・資本市場活性化に向けての提言」：
http://www.mof.go.jp/about_mof/councils/kinyuukaigou/report/gaiyou.pdf

³証券取引等監視委員会 HP「不正取引について」
<http://www.fsa.go.jp/sesc/support/hukousei/hukousei.htm>

⁴<http://hotcopper.com.au/>

た．このラベルは複数あるが，Delort らは”ramping ” (相場操縦の疑い) とラベル付けされた投稿を分析し，これらの投稿とリターン，ボラティリティ，出来高に相関があることがわかった．

このように，インターネットを用いた風説の流布の研究は行われてはいるが，実際に個人行動を分析した上で相場操縦士の可能性があるユーザの発見に関してはまだ研究は行われていない．

2.2 Web 上の犯罪捜査の研究

Web マイニングには，犯罪捜査に特化した Web Crime Mining と呼ばれる分野がある．この分野の研究については，[6] が関連研究をまとめている．

Web Crime Mining の手法は様々なものがあるが，その一つが異常検知である．この異常検知は得られたデータの中での外れ値を探すものである．この手法は，その通常時と異なる動きを検出する声質から，ネットワークへの侵入（アクセスの急激な増加）や，詐欺の検出に用いられてきた．しかし，ソーシャルメディアにおけるユーザの行動の異常度を発見することには用いられていない．

本研究の新規性をまとめると以下の通りである．

- 風説の流布を行うユーザの発見をテーマにした研究であること．
- ユーザ行動の指標に異常値を用いたこと，特に，複数ユーザの共起異常を分析したこと．

3 手法

3.1 実験手法

まず，本提案手法をまとめたシステムの概要は図 1 のようになっている．このシステムは，まず株価デー

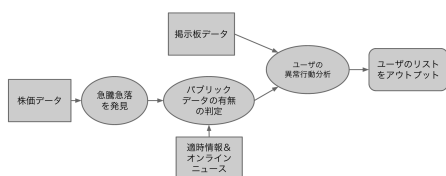


図 1: システムの概要図

タから急騰急落を発見し，その後パブリック情報と照らし合わせパブリック情報の影響を除去する．その後，急騰急落直前の書き込みを分析し，検出した異常行動にもとづいてユーザのランキングを作成するものであ

る．

それぞれの項目について，以下で説明する．

3.2 急騰急落の発見方法

今回の急騰急落の発見においては，忘却型自己回帰モデルを用いた計算アルゴリズムである，ChangeFinder アルゴリズムを用いた．ChangeFinder は，時系列データの各時点における外れ値スコア（本研究ではアノマリースコアと呼ぶ）を算出し，このスコアが高いほど，変化の割合が高いと判断する．

アノマリースコア算出のベースとなる考えは，自己回帰モデル（AR モデル）を用いた予測値と実測値との乖離の計測にある．スコアリングには AR モデルを用いており，ある時点 $t-1$ までのデータを学習し，時点 t での予測値と実測値の外れ値具合をアノマリースコアとして算出している．ここでアノマリースコアには対数損失を利用している．

このアルゴリズムの計算量はデータ数に対して線形オーダーに抑えられるため，4000 銘柄に関して一日あたり数百～数千の投稿を分析することに適していると言える．

また，実際山西らは金融データにおいてこのアルゴリズムを用い，急騰急落の発見において有効であることを示している [7]．

3.3 パブリックデータの導入について

今回取得したい急騰急落は株式掲示板の書き込みによるものであり，考えられる他の急騰急落の主要因であるパブリック情報については，基本的に株価はこれからの業績に対する期待値で動いているため，その影響を除外する必要がある．今回その除外方法として，ChangeFinder によって急騰急落を発見後，その 30 分前にその企業に関するパブリックデータがないかを検索した．

3.4 ユーザの異常行動の発見について

ユーザの行動は大別して 2 つあり，個別アカウントによるものと複数アカウントによるものである．

3.4.1 個別アカウントの異常行動の発見

今回はユーザが取れる行動として，書き込み行動，返信行動，評価行動の異常度を測定した．

書き込み頻度の異常

書き込み行動の異常度として、あるユーザが単位時間あたりに書き込む件数である書き込み頻度の異常度を測る。普段そこまで多く書き込んでいないユーザが、ある瞬間に急激に書き込んでいるとすれば、それは周りのユーザを煽ろうとする意思があると推察される。

その異常度の尺度としては、ポアソン分布を用いた。式は以下である。定数 λ に対し、自然数を値にとる確率変数 X が

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (1)$$

を満たすとき、確率変数 X はパラメータ λ のポアソン分布に従う。ここで、 e はネイピア数であり、 $k!$ は k の階乗、また λ は所与の区間で発生する事象の期待発生回数である。今回は、2014/12 2015/5 の場中の時間帯を 30 分スロットに区切り、その中のユーザの書き込み頻度を通常時の書き込み頻度のポアソン分布のモデルとした。ここで、 P の値が小さければ小さい程、それだけそのユーザの急騰急落前の書き込み行動が異常であることを示す。

返信行動、評価行動の異常

返信行動の異常に関しては、急騰急落前の書き込みの中で、どれくらい多く返信をしているかの異常度を測るものである。通常時に比べ、ある時間帯に多く返信行動をしてれば、その瞬間にできるだけ多くのユーザに働きかける意思があったものと思われる。返信行動異常の尺度に関しては、割合の偏りを測るのに一般的に使用されるカイ二乗値を用いた。急騰急落時の投稿数（返信以外）と返信数を C_{ar} , C_{or} 、通常時の投稿数（返信以外）と返信数を $C_{\hat{ar}}$, $C_{\hat{or}}$ とそれぞれ定義すると、このとき、 χ^2 値は次のように定義される。

$$\chi^2 = \frac{(C_{ar} - C_{or})^2}{C_{or}^2} + \frac{(C_{\hat{ar}} - C_{\hat{or}})^2}{C_{\hat{or}}^2} \quad (2)$$

この値が大きければ大きいほど、通常時に比べ多く返信行動をしているということである。

また同様の手法を用い、評価行動の異常度も測定した。

3.4.2 複数アカウントの異常行動の発見

掲示板では、個別のアカウントだけではなく、複数のアカウントによって相場操縦が行われる場合がある。それは、複数人による共謀行為であったり、掲示板の匿名性を活かしたなりすまし行為の場合もある。

複数アカウントの異常行動の指標には、共起関係の指標である Jaccard 係数を用いた。ここで書き込み頻度異常を求める際と同様に、場中を 30 分のスロットで

区切る。そして、あるユーザ u の書き込みスロット群を S_u とする。そのときユーザ u_1 と u_2 の Jaccard 係数の定義は式 3 のようになる。

$$Jaccard = \frac{|S_{u_1} \cap S_{u_2}|}{|S_{u_1} \cup S_{u_2}|} \quad (3)$$

この Jaccard 係数に関して、通常時 (o) と急騰急落時 (a) を比較するために、それぞれの Jaccard 係数を $Jaccard_a$, $Jaccard_o$ のように定義し、その割合をとることで、式 4 のように Suspicious_score を計算した。

$$Suspicious_score = \frac{Jaccard_a}{Jaccard_o} \quad (4)$$

ここで、Suspicious_score が高ければ高いほど、そのユーザ群が急騰急落時に多く同時に書き込んでいる可能性が高いことを示す。

3.4.3 ユーザネットワークの作成と分析

また、ここで取得されたユーザ群はあくまでペアのものであるが、実際の事件のように 3 つ以上のアカウントを使用している場合も想定される。そこで、この Suspicious_score で検出したペア群によってネットワークを作成し、そのネットワークを分析することで、3 人以上で同時に書き込んでいるユーザの検出を試みた。

4 使用したデータ

4.1 株価データ

実際に全ての銘柄において風説の流布が行われるわけではないため、本研究では特に風説の流布が行われやすい 320 銘柄において実験を行った。銘柄の選定条件は以下である。

- 発行株数が 5 千万株以下であること
- 資本金が 50 億円以下であること
- 株価が 300 円以下であること

この条件については [8] を参考にした。これらの銘柄について、分足のデータを東京証券取引所から購入し、使用した。

4.2 インターネット掲示板データ

本研究において使用した掲示板は、日本最大級の掲示板である textream である。

今回使用するデータは、2014/12～2015/5 までの半年間のデータであり、かつ 9:00～14:59 の場中のデータである（昼休みの時間帯を除く）。使用した 320 銘柄において、総ユーザ数 8,601、総コメント数 91,549 であった。

4.3 パブリック情報データ

表 1 に使用した記事の情報をまとめる。オンラインニュースに関しては、日本の市場の株価に影響を与える主要なメディアを選択した。

表 1: 使用したパブリックデータ

情報ソース	件数
適時開示情報	8,877
日経新聞	7,864
ブルームバーグ日本版	4,082
ロイター通信日本版	6,976
CNN 日本版	1,002
朝日新聞	3,992
読売新聞	5,211
産経新聞	14,560
ヤフーファイナンス	9,850
合計	62,396

表中の件数は、2014/12～2015/5 もので、かつ場中（9:00～14:59 除昼休み）に公開されたものである。

5 結果

本章では、前章で説明した実験手法によって実際に得られた結果について述べる。

5.1 書き込み頻度異常度分析の結果

書き込み頻度の異常分析においては、表 2 のような結果を得た。

表 2: 書き込み頻度の異常上位ユーザ

user	anomaly_slot	anomaly_comment	other_slot	other_comment	poisson_probability
not*****	1	10	2	2	7.60E-35
koj*****	1	7	9	15	1.34E-17
けん	1	8	24	46	1.67E-16
dea*****	2	13	4	11	1.11E-15
クロ	1	4	3	1	2.46E-14

この表は、左から

- ユーザのアカウント名 (user)
- ユーザが書き込んだ急騰急落の-slot数 (anomaly_slot)
- 急騰急落-slotにおいて書き込んだ書き込み数合計 (anomaly_comment)

- ユーザが書き込んだ通常時の-slot数 (other_slot)
- 通常-slotにおいて書き込んだ書き込み数合計 (other_slot)
- 通常書き込みをポアソン分布とした際の、急騰急落書き込み頻度の発生確率 (poisson_probability)

を示している。

一番右列のポアソン分布を用いた確率によって昇順にソートされている。この表を見ると、急騰急落前のみ異常な頻度で書き込んだユーザが取得できていることがわかる。例えばランキング 1 位のユーザは、通常時 2 slot に対して 2 回の書き込みというペースで書き込んでいるにも関わらず、急騰急落前には 30 分間で 10 回の書き込みを行っている。

実際にここで抽出されたユーザが、風説の流布と思われる書き込みをしているかどうかを確認するため、書き込み内容を確認する。

実際にこのユーザの急騰急落直前の書き込みを見てみると、表 3 のようなものが続く。発言の多くは「もっ

表 3: 書き込み頻度異常上位ユーザの急騰急落直前の書き込み

user	stock_ticker	date	comment_time	comment
not*****	9478	1/29/2015	14:28	掲示板を使った投資家へのマインドコントロールだとい。
koj*****	6819	2/5/2015	13:06	もっと買いまーす!
けん	6495	1/20/2015	14:29	【ヤバい】逃げときます! ヤバい感じなんで...
dea*****	3266	1/26/2015	13:26	そもそも下げる理由が見当たらない。他の新興銘柄と違って確固たる材料が出るしね。

と買いまーす」や「そもそも下げる理由が見当たらない。」など株価の評価や他人を煽るものが多く、これらの発言を見るだけでも他人に対して影響を及ぼそうとする様子が窺える。

5.2 返信行動異常度、評価行動異常度分析の結果

次に、銘柄に対しての返信行動に関する異常行動分析において得られた結果を示す。表 4 のようになっている。

表 4: 返信行動異常上位ユーザ

user	anomaly_reply_c	anomaly_comment_c	other_r_c	other_c_c	chi-square
drtetsu	22	49	158	542	565.49
tre*****	16	36	178	479	557.14
momoco	6	15	97	358	414.00
キャッシュハンタ 007	2	2	161	240	393.04
	4	12	65	252	285.82

この表は、左列から

- ユーザのアカウント名 (user)
- ユーザが書き込んだ急騰急落の返信付き書き込み数 (anomaly_reply_c)
- 急騰急落スロットにおいて書き込んだ書き込み数合計 (anomaly_comment_c)
- ユーザが書き込んだ通常時の返信付き書き込み数 (other_r_c)
- 通常スロットにおいて書き込んだ書き込み数合計 (other_slot)
- 通常書と急騰急落時を比較したカイ二乗値 (chi-square)

となっている。この表は一番右列のカイ二乗値で降順にソートされている。

この表を見ると、上位ユーザは急騰急落直前に通常時に比べ多く返信行動をしていることがわかる。例えば、このランキングで1位のユーザは、通常時542回の書き込みのうち158回が返信行動であるのに対し(約29%)、急騰急落直前では、49回の書き込みに対し、22回の返信行動(約45%)を行っている。

次に、銘柄に対しての評価行動に関する異常行動分析において得られた結果を示す。表5のようになっている。

表 5: 評価行動異常上位ユーザ

user	anomaly_senti_c	anomaly_comment_c	other_s_c	other_c_c	chi_square
xko*****	3	3	115	132	235.15
star_holder	9	9	129	130	224.25
元気	4	4	63	75	122.47
わごん	2	3	18	112	120.30
漫 株太郎	5	5	43	64	87.97

この表の列の項目は、返信行動の項目のうち「返信」を「評価」に置き換えたものである。この表を見ると、上位ユーザは急騰急落直前に通常時に比べ多く評価行動をしていることがわかる。例えば、このランキングで1位のユーザは、通常時132回の書き込みのうち115回が評価行動であるのに対し(約87%)、急騰急落直前では、3回の書き込みに対し、3回の評価行動(100%)を行っている。

5.3 複数アカウントの異常行動

次に、複数アカウントの異常行動に関して分析結果について述べる。

表 6: 共起関係異常度

user1	user2	a_jaccard	o_jaccard	Suspicious_score
pro*****	sup*****	0.125	0.011	11.70
kur*****	ibu*****	0.333	0.029	11.47
pro*****	チャネラ	0.091	0.009	9.82
v*****	トム12	0.333	0.036	9.17
ibu*****	sup*****	0.167	0.022	7.60

5.3.1 Jaccard 係数を用いた共起関係異常度の分析

表6は、ユーザの共起関係異常度を分析して得られた結果の表である。

表6は、左列から、

- ユーザ1のアカウント名 (user1)
- ユーザ2のアカウント名 (user2)
- 急騰急落時のユーザ1とユーザ2のJaccard係数 (a_jaccard)
- 通常時のユーザ1とユーザ2のJaccard係数 (o_jaccard)
- 急騰急落時と通常時のJaccard係数の比率 (Suspicious_score)

という項目になっている。ここで、ランキング上位のユーザペアを見てみると、急騰急落時のJaccard係数が0.125であるのに対し、通常時は0.011しかなく、通常時に比べ、急騰急落直前に共起関係が非常に強くなっていることがわかる。

このランキングによって、急騰急落時のみに異常に同タイミングで書き込んでいるユーザペアが検出できた。

5.3.2 ネットワーク分析を用いた相場操縦グループの分析

ここでは、共起行動によって抽出されたユーザのペアリストから、ユーザの共起グループの発見を試みる。図2が、今回得られた分析によって得られたネットワーク図である。

ここで、各ノードはユーザを示す。ノードの大きさはanomaly_slotの大きさを示している。エッジは今回得られたSuspicious_scoreを示しており、太いほど値が高い。この図により、異常な共起関係をとっていたユーザがグループを作成していることがわかった。

6 評価

6.1 システムの評価手法

今回は有識者として金融の業務経験者の方々にアンケートをとることで、風説の流布と思われる書き込み

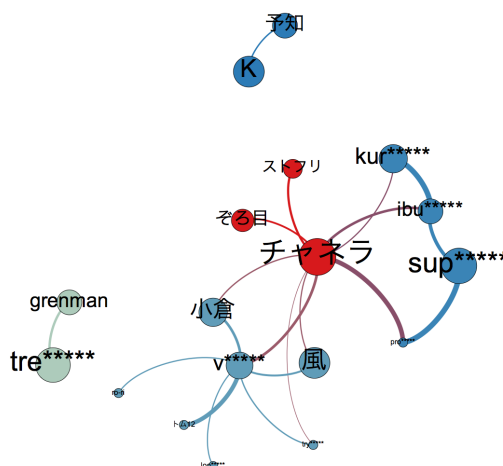


図 2: Jaccard 係数の異常度により作成されたネットワーク図

の評価を行った。

評価を行う分析は、本手法のベースになっている「書き込み頻度の異常」である。以下のタイプ別の書き込み 100 件ずつの 4 つの群から、一番「風説の流布」と思われる書き込みの集合であると思われるものから 1 ~ 4 位の順位をつけてもらうことにする。

- リスト上位ユーザの急騰急落前の書き込み (D)
- リスト上位ユーザの通常時の書き込み (B)
- 非リスト上位ユーザの急騰急落前の書き込み (C)
- 非リスト上位ユーザの通常時の書き込み (A)

6.2 結果

アンケートの結果を 7 に示す。

表 7: アンケート結果の集計表

書き込み群	リスト上位ユーザ	急騰急落前	1位	2位	3位	4位
A			5	1	0	4
B			1	2	4	3
C			0	5	4	1
D			4	2	2	2

表 7 はアンケートの集計結果である。

「リスト」上位ユーザは今回の手法によって取得できたユーザの書き込みを示している。

「1 位」~「4 位」は回答者が答えた回答である。最も風説の流布を取得できていると思われる書き込み群を 1 位に選んでもらっている。

「急騰急落前」は株価の急激な変動前の書き込みかどうかをしめしている。

この表からわかることを以下にまとめる。

- 提案手法で取得できた D が安定的に上位を獲得できている。
- B, C はそれぞれ中程度の順位を多く取得できている。
- A は 1 位取得回数は 1 位だが、4 位取得回数も 1 位である。

この結果から、D が最も上位の結果を示しているため、基本的には本提案手法で意図した結果が取得できていることがわかった。本手法は発言の異常度からユーザをランキング化するものであるが、その上位のユーザのコメントのみを他人に見せたことでこのような結果を得たことは、本手法の有効性を示していると言える。

7 考察

7.1 手法について

元々急騰急落という株価の動きを追うような研究は今までなかったが、今回このアルゴリズムを使用することで、適切に株価の「急騰急落」というパターンを使用できたとと思われる。

また、パブリックデータの導入についても、適時開示情報については、上手く取得できているように思える。しかし一方で、オンラインニュースの取得については、取得できたものが殆ど経営に関わるようなニュースではなく、単なる銘柄の株価情報であったため、上手く取得できたとは言えないと思われる。

原因としては、まず企業の検索方法が上手くいっていないことが挙げられる。例えば同じ株式会社でも (株) や、株式会社、KK など、その表記は媒体によって異なる。これらの精度を上げることは今後の課題であると思われる。

また、他のニュースにおいて企業名が直接言及されなくても、影響を与えているケースを考えることも課題である。

7.2 実験結果について

評価、返信、書き込みに関しては、それぞれ上手く取得できていると思われる。今後の課題があるとすれば、書き込み内容の異常検知が考えられる。別書き込みを一つ一つ分析するような場合はなおさらである。

複数アカウントの行動においては、今回通常の Jaccard 係数ではなく、通常時と急騰急落時を比較した Jaccard 係数の異常度を使用した。目的として「普段同時に書き込まないユーザ群が一緒に書き込む」という現象を取る上では理にかなっていると言える。実際のネット

ワークについては、単なるペアの集まりではなく、実際に3人異常のユーザによるネットワークが形成され、それを取得できた点に関して、手法は効果的だったと言える。より深い分析のため、リプライ関係や書き込み内容の類似度などを分析するとより精度が高くなると思われる。

7.3 評価について

アンケートによって本手法の一定の有効性を示すことはできたが、アンケート回答者をもっと多くすることも課題であると思われる。また、実際の現場で使用してみて、その有効性を評価されることも今後の展望として望まれる。

8 まとめ

本研究では株式掲示板において、ユーザの行動異常度を測定することにより、相場操縦行為を発見する手法について提案した。提案手法は、株価の急騰急落を前における掲示板でのユーザの行動を、様々な側面から分析することにより、ユーザの異常行動具合をランキング化するものである。

その結果、実際に風説の流布を試みた可能性のあるユーザのリストを作成することができた。

さらに、そこで得たランキング上位のユーザの急騰急落前の書き込みに対し、アンケートによって評価することで、本提案手法の有効性が評価できた。

本研究で提案された手法が実際の相場操縦発見の現場で使用されることになれば、著者の望むところである。

参考文献

- [1] Thomas C Chiang and Dazhi Zheng. An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, Vol. 34, No. 8, pp. 1911–1921, 2010.
- [2] Félix Villatoro. The delegated portfolio management problem: reputation and herding. *Journal of Banking & Finance*, Vol. 33, No. 11, pp. 2062–2069, 2009.
- [3] Jean-Pierre Zigrand. Endogenous market integration, manipulation and limits to arbitrage. *Journal of Mathematical Economics*, Vol. 42, No. 3, pp. 301–314, 2006.
- [4] David J Leinweber and Ananth N Madhavan. Three hundred years of stock market manipulations. *The Journal of Investing*, Vol. 10, No. 2, pp. 7–16, 2001.
- [5] Jean-Yves Delort, Bavani Arunasalam, Maria Milosavljevic, and Henry Leung. The impact of manipulation in internet stock message boards. *International Journal of Banking and Finance*, *Forthcoming*, 2009.
- [6] Javad Hosseinkhani, Mohammad Koochakzaei, Solmaz Keikhaee, and Javid Hosseinkhani Naniz. Detecting suspicion information on the web using crime data mining techniques. *International Journal of Advanced Computer Science and Information Technology*, Vol. 3, No. 1, pp. 32–41, 2014.
- [7] Kenji Yamanishi and Jun-ichi Takeuchi. A unifying framework for detecting outliers and change points from non-stationary time series data. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 676–681. ACM, 2002.
- [8] 中原圭介. 仕手株でしっかり設ける投資術. 日本実業家出版, 2005.

Twitter 位置情報・テキスト情報を用いた人の移動モデル構築と 観光地推薦手法の提案

Proposal for Tourist Attractions Recommendation and Mobility Model Based on Twitter's Spatio-temporal Data and Text Data

前田 高志ニコラス^{1*} 吉田 光男² 鳥海 不二夫¹ 大橋 弘忠¹

Takashi Nicholas MAEDA¹, Mitsuo YOSHIDA², Fujio TORIUMI¹ and Hirotada OHASHI¹

¹ 東京大学大学院工学系研究科

¹ Graduate School of Engineering, The University of Tokyo

² 豊橋技術科学大学

² Toyohashi University of Technology

Abstract: This research investigates a method to show important tourist attractions' locations and evaluate each attraction's attractiveness by using fine-grained spatio-temporal data of Twitter. We modify the model called Gravity Model known in human geography and apply it to estimate each attraction's attractiveness based on the numbers and the distances of human trips extracted from Twitter data. We distinguish whether each tourist attraction is just a convenient place or a place which has an area-specific attractiveness according to each location's distribution of TF-IDF.

1 はじめに

地方経済の疲弊や高齢化、震災後の復興に向けた動きの中で、地方都市の存立にとって、人の活発な地域間移動が重要であるという認識が広まってきた。中でも観光目的による人の流入は地方経済の重要な要素であり続けている。一方、観光庁などにより Twitter 位置情報を利用した観光情報抽出に注目が集まっており、Twitter 位置情報を用いた研究も増加している。国内の観光地の把握は観光業界にとっても各自治体や日本政府にとっても重要課題であり、また、外国人への情報提供という点でも喫緊の課題である。

魅力のある観光地情報の抽出のためには 2 つの分析すべき点があると考えられる。一点目は、実際に人がどこからどこへどれだけ移動しているのかという、移動量・移動距離の傾向の分析である。二点目は、移動の目的が観光目的であるかを峻別することである。

これまで、人の住んでいる場所と移動先の関係を考慮したものとして、人文地理学の領域では、人や物や情報の空間的フローを説明する空間的相互作用モデルの構築がなされてきた [2]。特に歴史が深いものはニュートンの万有引力を用いた重力モデルである。これは、2 都市間の流量が両都市の規模の積に比例し、距離に反比例するとしたモデルである。空間情報学の領域では、

Phithakkitnukoon ら [8] が、大規模な携帯電話の位置情報をもとにして、旅行者の行動を詳細に分析することに成功している。この研究では、旅行者の旅行頻度、移動距離、目的地、出発地、移動手段、現地での滞在時間の関係性を個人レベルで抽出し、その傾向の分析を行った。一方、Twitter を利用した移動研究としては、観光ではなく日々の移動に注目した研究として、若宮ら [6] が、位置情報付き Twitter データとパーソナリティ調査のデータをもとに、群衆の移動傾向を分析している。この研究では、地域間の移動について、移動距離・移動時間・移動量の 3 つの値をもとに各地域間の直感的な近接性を多次元尺度構成法により示すことに成功している。しかし、これまでに述べたように様々な先行研究はあるものの、Twitter 位置情報を利用した住所と観光地に関する広域な研究はなされていない。

移動目的については、Yin ら [4] や Kurashima ら [7] の研究のような、ジオトピックモデルという分野が注目を集めている。これらは、文書における語の共起傾向から潜在トピックを抽出するトピックモデリングという技術を、ジオタグを持つソーシャル・メディアへ応用し、各地域の持つ特徴を理解するというものである。しかし、これによって「スポーツ」や「アート」や「エンタテインメント」といった地域の特徴分類が可能になったとしても、各々の場所が「観光地」であるか否かの判断は難しいのではないかと考えられる。なぜな

*連絡先：東京大学大学院工学系研究科
〒 113-8656 東京都文京区本郷 7-3-1
E-mail: maeda@crimson.q.t.u-tokyo.ac.jp

ら、通常であれば観光目的でない他の機能を持つ場所が観光地として機能することがあるためである。例えば、アメリカのゴールデン・ゲート・ブリッジは橋であるが、一大観光地である。また、一部の名門大学は観光目的で人が訪れる場所もある。このようなものを横断して「観光地」として語の共起傾向を抽出するのは難しいと考えられる。したがって、観光目的の移動であることを判別する別の方法を考案する必要がある。

本研究では位置付き Twitter データから得られる地域間移動情報から人の移動モデルを構築し、「各地の魅力」「距離が移動件数に与える影響」を算出する。また、Twitter のテキスト情報により各地域の魅力の原因がその地域固有の魅力に基づくものなのか、利便性に基づくものなのかを判定する。最終的にこの情報をもとに旅行者への、住所に応じた観光地推薦技術に活かす。

本研究は以下の手順によって前記の目的を達する。

1. まず、Twitter データから地域間移動数を抽出するため、各ユーザの居住地と旅行・おでかけ目的の移動先を推定し、全ユーザについて集計する
2. 地域間移動数と目的地の魅力の間に成り立つモデルを立て、地域の持つ魅力を算出する
3. テキスト情報から、「他の地域にない魅力」を持つ場所を推定し、単に「便利な場所」と区別することで観光地を抽出する

本稿では最終的に、日本、イギリス、アメリカのデータを用いて、上記手順により観光地抽出を行った。

2 各ユーザの居住地と移動先の分類

人の移動はもっぱら自宅や職場、学校といった特定の地点を日々往復する移動に占められている。Esterら [3] の DBSCAN(Density Based Spatial Clustering Algorithm with Noise) を用いた、人の重要地点の探索方法が複数考案されている。本研究では、DBSCAN を Twitter 分析用に改良したアルゴリズムによって、各ユーザの居住地と移動先を特定する。

1. 一人のユーザに関し、期間内の全 Tweet の位置を抽出する
2. その中からどれか 1 点を取り出し、半径 4km 以内に 4 日以上以上の Tweet がなければ、その点を Noise とみなす
3. 4 日以上以上の Tweet があれば、それらを同じクラスタとみなす

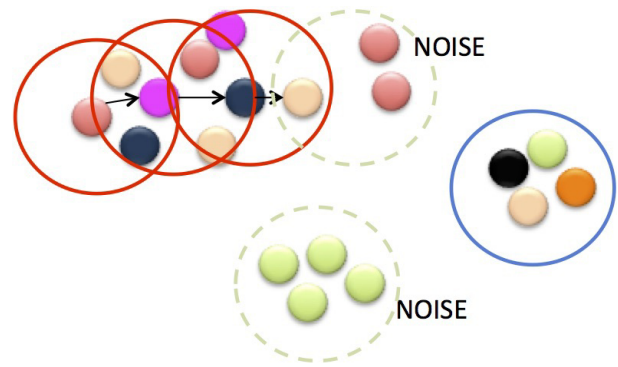


図 1: DBSCAN(異なる色は異なる日の Tweet であることを示す)

4. クラスタ内の別の点と同様に半径 4km 以内に 4 日以上以上の点を含むなら、それらの点も同じクラスタとする
5. クラスタを形成する点は高頻度滞在地とし、Noise となった点は低頻度滞在地とする

このようにして図 1 のように、各点をクラスタと Noise に分類することができる。ここで、最も多くの日数の Tweet を持つクラスタの重心をそのユーザの居住地とし、そこから低頻度移動地を結んだものをそのユーザの地域間移動とする。4 日をしきい値としたのは、国土交通省観光庁の観光白書 [10] において、日本人の 1 回あたりの旅行の宿泊日数が 2.1 泊とあることから、4 日間同じ場所に滞在する旅行は稀であるという考えに基づいて設定した。距離に関するしきい値 4km に関しては暫定的な設定であり、期間のしきい値も含め、最適なしきい値の獲得及び検証は今後の課題とする。

3 出発地・到着地のクラスタリング

前述の手順で数多くの出発地と到着地が得られるが、図 2 のように空間的に近いものをひとつにまとめた。そこで Mean Shift Clustering [1] を用いる。Mean Shift Clustering とは以下のような、漸次的な手順によって近い点同士をひとつにまとめる手法である。各ステップごとに各点が次に移る先の点を、自身を含めた近傍半径 r_n の円内に含まれるすべての点の重心とする。すべての点について、近傍半径 r_n の円内の点がそれよりさらに小さい収束半径 r_c の円内に収まれば、そこでこの処理を終了する。同じ収束半径に収まった点同士を同じクラスタとしてまとめ、各クラスタの重心をそのクラスタの代表点とする。

本研究では、出発地のクラスタリングについては、 r_n を 5km、 r_c を 1km とし、到着地のクラスタリングについては、 r_n を 2km、 r_c を 1km とした。また、毎ス

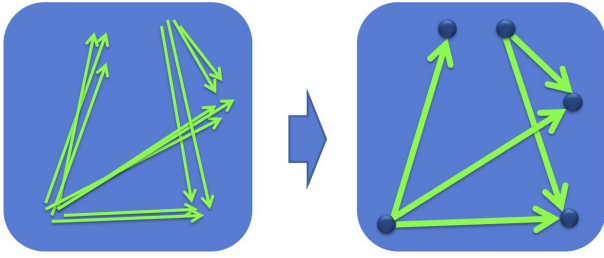


図 2: 出発地・到着地のクラスタリング

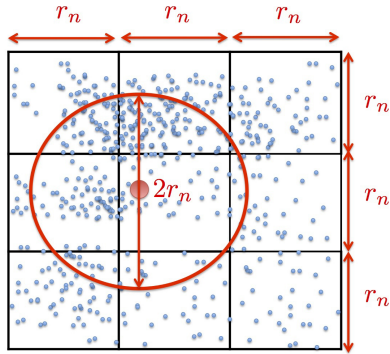


図 3: Mean Shift Clustering の効率化

トップですべての点同士の距離を求めると計算時間が膨大になるため、不要な計算を省く工夫をする。緯度・経度を等間隔で区切ったメッシュを張り、それらの各メッシュの縦・横の長さが r_n よりも長くなるようにする。このようにすれば、各点について、図 3 のようにその点が含まれるメッシュと近傍の 8 メッシュ内の点以外は r_n 以上の距離にあるため、距離の計算をせずすむ。これによって計算時間の短縮をはかる。なお、経度 1 秒あたりの長さは赤道から離れるほど短くなるため、分析対象国の最北端の緯線における経度が r_n となる単位でメッシュを形成する。

4 モデル化と指標値の定義

4.1 概要

本研究では、2 地点間の移動件数は下記 4 つの値によって決まると考え、到着地の魅力を推定するために、これらの指標値が互い持つ関係性をモデル化する。

- 出発地の放出力
出発地の放出力が高ければ、その地点からの移動件数と移動距離が大きくなる。
- 到着地の魅力
到着地の魅力が高ければ、その地点への移動件数と移動距離が大きくなる。

- 移動コスト

2 地点間の移動コストが高ければ、その区間の移動件数が減る。本研究では、移動距離をコストの指標値として用いる。

- 到着地の競合

出発地の周りに多くの魅力ある到着地があれば、それぞれの到着地への移動件数が分散する。

4.2 目的地選択のモデル

出発地点 s に存在するユーザが、数ある目的地から目的地 e を選択する確率 $P(s \rightarrow e|s)$ を以下の式で表す。

$$P(s \rightarrow e|s) = \frac{A_e}{D_{se}^\alpha} / E_s \quad (1)$$

- A_e : 目的地 e の持つ絶対的な魅力 (未知変数)
- D_{se} : 出発地 s と到着地 e の間の距離 (既知変数)
- α : 距離が目的地の魅力に与える影響を決定する係数 (未知変数)
- E_s : 出発地 s の周囲の魅力の総和 (未知変数)

$$E_s = \sum_k \frac{A_k}{D_{sk}^\alpha} \text{ によって与える。}$$

出発地点 s に存在するユーザが、数ある目的地から目的地 e を選択する確率 P は、実データによる観測値を用いると下記のように表すことができる。

$$P_{obs}(s \rightarrow e|s) = \frac{T_{s \rightarrow e}}{\sum_k T_{s \rightarrow k}} \quad (2)$$

- $T_{s \rightarrow e}$: データで得られた、出発地 s から目的地 e へ移動した移動件数 (既知変数)

4.3 重回帰分析による指標値の算出方法

式 (1) と式 (2) を等号で表し、両辺の対数を取り、重回帰分析が可能ないように式変形を行う。

$$\log \left(\frac{T_{s \rightarrow e}}{\sum_k T_{s \rightarrow k}} \right) = \sum_i x_i \cdot \log A_i - \alpha \cdot \log D_{se} - \sum_i y_i \cdot \log E_i + c \quad (3)$$

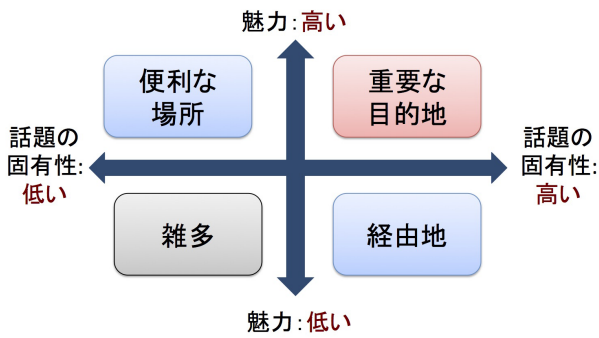


図 4: 地域の魅力と固有性

ここで、左辺は被説明変数であり、データから求めることができる。説明変数は x_i , D_{se} , y_i であり、係数は $\log A_i$, α , $\log E_i$ である。また、 c は定数項である。説明変数の D_{sd} は緯度経度から GRS80 楕円体をもとにした計算によって求める。また、 x_i は入力データの目的地が e であるとき、 $i = e$ である x_i を 1 とし、それ以外を 0 とする。同様に y_i は入力データの出発地が s であるとき、 $i = s$ である y_i を 1 とし、それ以外を 0 とする。

Mean Shift Clustering をもとに出発地・到着地をクラスタリングしたのちに、各 2 地点間の移動件数を算出し、式 (3) に代入して、重回帰分析により、各目的地の魅力 (A_e)、距離が魅力に与える係数 (α)、各出発地の周囲の魅力 (E_s) を求める。

5 テキスト情報による到着地の固有性判定

5.1 地域の魅力と話題の固有性

移動モデルによって各地の魅力を求めることが可能であるが、その魅力がその土地の持つ固有性に起因するのか、あるいは単純に利便性があるからだけなのかを判断する必要がある。ここでは、Twitter のテキスト情報を用いて、ユーザたちの Tweet からその土地の固有性に着目した投稿が多いのか、あるいは、ほかの多くの地域でも話される話題が多く占めているのかに注目することで、固有性の判定を行う。これにより、図 4 の第 1 象限のように魅力が高く話題の固有性も高い場所は、その土地の持つ固有の魅力に人が惹かれて訪れていることを示す。魅力が高く話題の固有性が低い場所は、大規模店舗などの利便性の高い場所を示すと考えられる。魅力が低く、話題の固有性が高い場所は経由地として一時滞在する場所を示すと考えられる。

5.2 テキスト情報に注目した固有性の算出

土地の固有性を評価するために、Twitter 投稿記事のテキスト情報を用いる。各到着地において投稿された Tweet のうち、各ユーザの低頻度移動地のもを集め、それをまとめてひとつの文書とする。文書内での出現回数が多い単語が、他の文書でほとんど出現しなければ、その地域は他にはない魅力を持っていると考えられる。例えば、富士山の周辺では「富士山」「御来光」「〜合目」という単語が多く、これは他の場所でほとんど出現しない。逆に「ご飯」「テレビ」「サッカー」といったありふれた単語が主要となる地域では、固有の魅力をあまり持たないと考えられる。

関連研究として、ジオタグの付与されていない Twitter 投稿記事のテキスト情報から投稿時の地理的位置を推定する Cheng ら [5] の研究がある。これらの研究では単語の地理的な局所性に注目して投稿時の地理的位置を推定している。三木ら [9] は、単語の地理的局所性を算出するために、ジオタグのついた投稿記事から各場所ごとの単語の TF-IDF を算出している。

本稿は地理的局所性の高いローカル語が投稿される割合が多い場所とそうでない場所を判定することで、その土地の固有性を評価する。ここでは、三木ら [9] と同様に、キーワードの重要度を表現する TF-IDF を指標に用いる。これは主に文書の特徴づける単語に高い値を割り振るものであり、文書の要約や同ジャンルの文書のクラスタリングに用いられる。TF-IDF は TF (Term Frequency: 単語の出現頻度) と IDF (Inverse Document Frequency: 逆文書頻度) の積によって求める。

$$tfidf_{w_i,d} = tf_{w_i,d} \cdot idf_{w_i,d} \quad (4)$$

$$tf_{w_i,d} = \frac{N_{w_i,d}}{\sum_k N_{w_k,d}} \quad (5)$$

$$idf_{w_i,d} = \log \frac{|D|}{|d : d \ni w_i|} \quad (6)$$

ここで、 $N_{w_i,d}$ は文書 d に含まれる単語 w_i の出現回数を示す。 $|D|$ は全文書数を示す。

地域の固有性が高い場所では、ユーザが他の地域にない、その地域固有の単語を多く発し、地域の固有性が低い場所ではその逆となると考えられる。このため、前者の地域の文書（その地域の全 tweet のテキスト情報の結合）では、一部の単語の TF-IDF が極端に高くなり、その他大部分の単語との落差が大きくなる。逆に地域の固有性が低い場所では、文書内の TF-IDF の分布はよりなだらかなものとなる。このためここでは、各文書の TF-IDF が高いものから上位 10 % のものの総和をその地域の固有性と定義づける。

表 1: 有効ユーザ数と決定係数

	期間	有効ユーザ数	決定係数
日本	2014/4	81115	0.75412
	2014/5	79870	0.75431
	2014/6	86167	0.77979
	2014/7	93809	0.77449
	2014/8	107418	0.72844
	2014/9	95723	0.73748
	2014/10	85012	0.75326
	2014/11	83743	0.74258
	2014/12	106951	0.74248
	2015/1	105444	0.74380
	2015/2	99846	0.76183
	2015/3	124954	0.73715
	イギリス	2014/8	164040
2014/9		146360	0.74169
2014/10		115143	0.72238
アメリカ	2014/8	716413	0.64896

表 2: 距離が移動件数に与える影響

	期間	距離の累乗係数
日本	2014/4	0.93051
	2014/5	0.94589
	2014/6	0.91594
	2014/7	0.90771
	2014/8	0.86430
	2014/9	0.94081
	2014/10	0.91642
	2014/11	0.94743
	2014/12	0.91262
	2015/1	0.85424
イギリス	2014/8	0.69934
	2014/9	0.81700
	2014/10	0.87660
アメリカ	2014/8	0.68317

6 実データによる計算結果

6.1 データセットと指標値の計算結果

日本のデータについては、2014年4月から2015年3月にかけて、1ヶ月ごとに計算を行った。イギリスのデータは2014年8月から10月にかけて、1ヶ月ごとの計算を行い、アメリカ（アラスカとハワイを除く48州）のデータについては、2014年8月の1ヶ月分のみのデータで計算を行った。ただし、アメリカについては地域数が膨大であったため、メモリの制約から、100以上の目的地への移動を持つ出発地と、100以上の出発地からの移動を持つ目的地以外を計算の対象から除外した。ただし、このような除外を行うと式(3)の重回帰分析の決定係数が低くなる傾向があったため、今後このメモリ制約の問題について解決する必要がある。

表1は各月の有効ユーザ数と式(3)の重回帰分析の決定係数を表す。ここで有効ユーザ数とはDBSCANによって、高頻度滞在地と低頻度滞在地の両方を持つユーザの数を示す。式(1)の距離が目的地の魅力に与える影響を決定する係数 α は表2の通りであった。すべて0.68~0.95の範囲に収まった。これが指し示すのは、出発地から近距離にある到着地は距離が増えるに伴って大きくその魅力を減じるが、遠距離にある到着地同士を比べる場合、到着地の魅力への距離の影響は比較的緩やかになることを意味する。

6.2 魅力と固有性に基づく地域クラスタリング結果

次に、日本・イギリス・アメリカの2014年8月分のデータについて、移動モデルに基づいた地域の魅力と、テキスト情報に基づいた話題の地域固有性の値を用いて、地域のクラスタリングを行った。クラスタリングの方法は、魅力、固有性のそれぞれについて最大値で値を割ったものを使用して、2次元空間上に位置づけ、それをk-means法を用いて4つのクラスタに分割した。その結果を図5~10、表3に示す。

日本においては、各クラスタの分類は図5の通りのもとなった。魅力と固有性の両者が高いクラスタ1はディズニーランド、ユニバーサル・スタジオ・ジャパン、沖縄美ら海水族館、京都伏見稲荷大社などの観光地が多く含まれたほか、この期間内に行われたロック・フェスティバルの地域が含まれていた。魅力が高く、固有性の低いクラスタ2はイオンを含む、地方の大型商業施設、ショッピングモール、大規模店舗が多く含まれていた。魅力が低く、固有性の高いクラスタ3はサービスエリア・パーキングエリアのような高速道路の中継地点が多く含まれており、地図上でも高速道路沿いに多かった。さらに、東京湾の対岸にある2つのフェリー乗り場（久里浜と金谷）もこのクラスタに含まれていた。

イギリス・アメリカにおいては、図7・9の通り、クラスタ1は日本と同様に魅力も固有性も高い観光地クラスタとなったが、その他のクラスタについては、固有性の違いはなく、魅力の大小によってのみ別れるこ

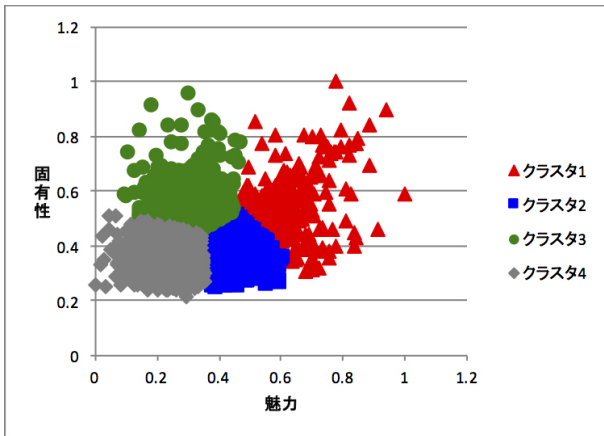


図 5: 【日本】魅力と固有性に基づく地域のクラスタリング (2014年8月)

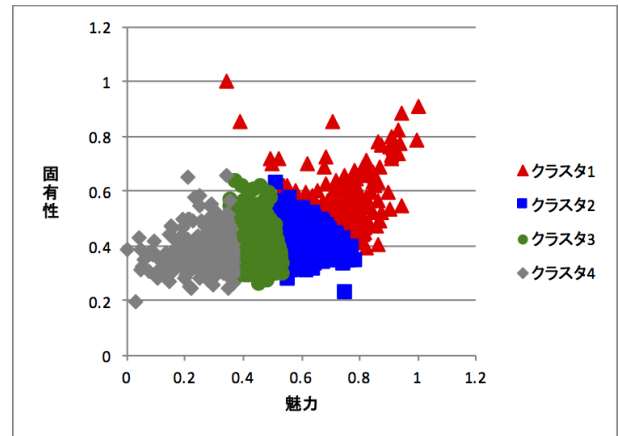


図 7: 【イギリス】魅力と固有性に基づく地域のクラスタリング (2014年8月)

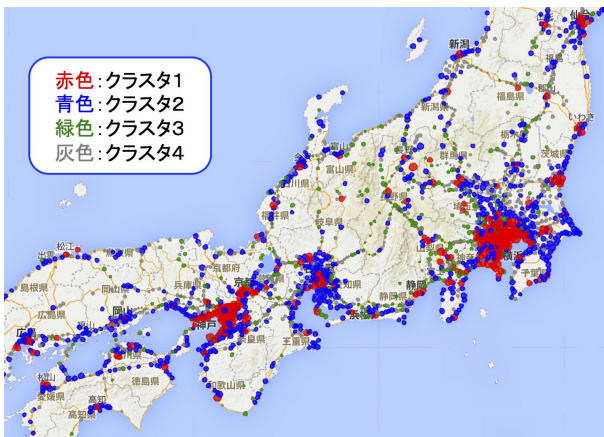


図 6: 【日本】各クラスタの地図上の分布 (2014年8月)

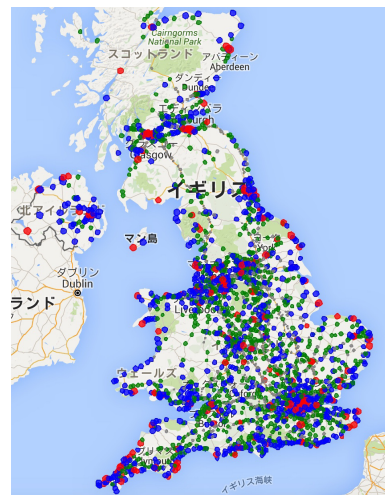


図 8: 【イギリス】各クラスタの地図上の分布 (2014年8月)

とになった。イギリスにおいてクラスタ1に分類されたのは、大都市の中心地、サッカースタジアム、ロックフェスティバル、古城、ストーンヘンジのような遺跡、湖水地方などであった。アメリカにおいては、遊園地、ビーチ、カジノ、アメリカンフットボール・スタジアムなどであった。ゴールデン・ゲート・ブリッジについてもクラスタ1に分類された。

以上より、いずれの国においても、魅力と固有性がともに高いクラスタには、ジャンルを問わず広い範囲で観光地が抽出された。さらにゴールデン・ゲート・ブリッジのように、第一義的な目的が観光以外の用途でありながら観光地として機能する場所も抽出することにも成功した。また、いずれの国でも魅力が高く固有性が低い場所は利便性の高い場所が分類された。さらに、日本においては固有性が高く、魅力が低いクラスタが抽出でき、ここでは滞在時間が短く、特定機能のみを持つ経由地が多く含まれていた。イギリス・アメリカでこのようなクラスタが抽出できなかった原因を

明らかにすることは今後の課題である。現時点で考えられるのは、クラスタリング手法の問題であるという点と、交通環境の違いから生じたという点が考えられる。前者は、k-means以外のクラスタリング手法を用いることや、あるいは分布の重心を求めて、その点から縦軸と横軸を引いて分割する手法を取れば同様のクラスタを抽出できる可能性があるというものである。後者は、日本においては山岳地帯（非可住地）が多く、その間を道路が走っているために、サービスエリアのような経由地が多数存在するのに対し、イギリスやアメリカでは平野部（可住地）が多く、日本のように経由地があまりない、というものである。今後、この点については明らかにしていく。

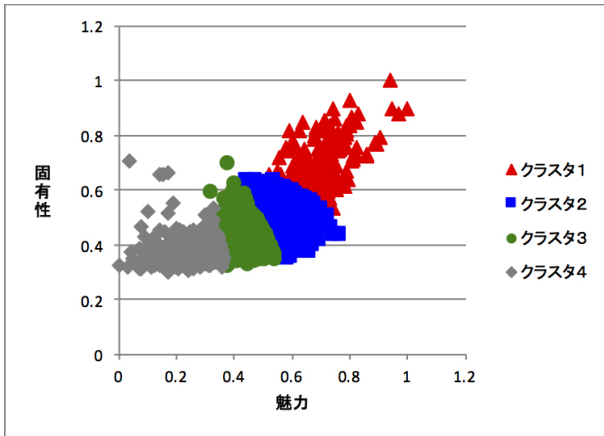


図 9: 【アメリカ】魅力と固有性に基づく地域のクラスタリング (2014 年 8 月)

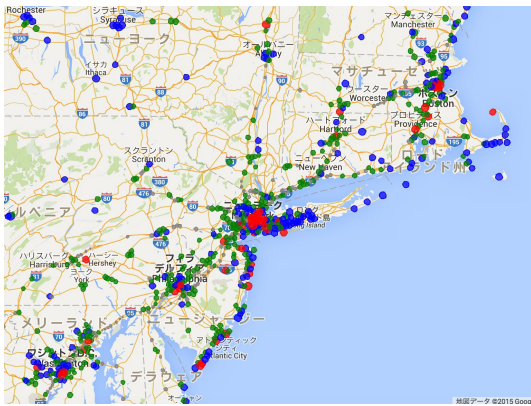


図 10: 【アメリカ】各クラスタの地図上の分布 (2014 年 8 月)

7 おわりに

本研究では、Twitter 位置情報を用いて、人の高頻度滞在地と低頻度移動地を求め、移動件数と距離から地域の魅力を算出した。また、テキスト情報から地域の固有性を算出した。そして、地域の魅力と固有性の値を用いて、地域を 4 つのクラスタに分けた。その結果、魅力の高い地域の中から、その地域が持つ固有性によって魅力が高まっている場所を抽出し、利便性によって魅力が高まっている場所と区別することができた。これにより、幅広いジャンルを横断して観光地を抽出することができた。

今後はジオトピックモデルの手法をともに用いることで、各地域がどのような条件のもとであれば魅力や固有性が高まるのかを検証する。また、今回導入した魅力と固有性が、ユーザへの観光地推薦役立つことを示すために、Kurashima らの研究 [7] における、協調フィルタリングに組み込むことで性能が高まるか否か

表 3: 各クラスタが含む地域数 (2014 年 8 月)

	日本	イギリス	アメリカ
クラスタ 1	353	192	216
クラスタ 2	1054	812	1060
クラスタ 3	354	955	1205
クラスタ 4	890	478	473

を検証する。ここで大幅に性能が上がれば、魅力と固有性が人の移動に大きく影響を与えることを実証できる。最終的にはこの視点を使うことで、豊かな都市計画に活かせることになるようにしたい。

参考文献

- [1] Fukunaga, K., Hostetler, L.: The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Transactions on Information Theory, 1975
- [2] 石川 義孝: 空間的相互作用モデル—その系譜と体系, 地人書房, 1988
- [3] Ester, M., Kriegel, H., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 1996
- [4] Yin, Z., Cao, L., Han, J., Zhai, C., Huang, T.: Geographical Topic Discovery and Comparison, in: Proceedings of the 20th International World Wide Web Conference, 2011
- [5] Cheng, Z., Caverlee, J., Lee, K.: You are where you tweet: a content-based approach to geo-locating twitter users, in: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, 2010
- [6] 若宮 翔子, 李 龍, 角谷 和俊: 位置ベース SNS を通じた群衆の移動経験に基づく都市空間の近接性分析, 情報処理学会論文誌, 2013
- [7] Kurashima, T., Iwata, T., Hoshide, T., Takaya, N., Fujimura, K.: Geo Topic Model: Joint Modeling of User's Activity Area and Interests for Location Recommendation, in: Proceedings of the 6th ACM International Conference on Web Search and Data Mining, 2013
- [8] Phithakkitnukoon, S., Teerayut Horanont, T., Witayangkurn, A., Siri, R., Sekimoto, Y., Shibasaki, R.: Understanding tourist behavior using large-scale mobile sensing approach: A case study of mobile phone users in Japan, Pervasive and Mobile Computing, 2014
- [9] 三木 翔平, 新田 直子, 馬場口 登: 単語の地理的局所性の経時変化を考慮したツイートの発信位置推定, 第 6 回データ工学と情報マネジメントに関するフォーラム, 2014

[10] 国土交通省観光庁: 平成 27 年版観光白書,
<http://www.mlit.go.jp/common/001095743.pdf>,2015

11月12日(土) 発表分の受賞理由

10. SIG-ALST 先進的学習科学と工学研究会

「Can LOD Based Question Generation Support Work in a Learning Environment for History Learning?」

Jouault Corentin, Kazuhisa Seta, Yuki Hayashi (Osaka Prefecture University)

(SIG-ALST-B503-07, 第 76 回, 2016 年 3 月 6 日)

選考理由:

本論文は、LOD (Linked Open Data) を用いた自己学習支援に関する研究である。LOD の普及拡大に伴い、学習者はLOD を通じて様々な知識源にアクセスできるようになったが、情報過多のため学習の目的を見失ったり、適切な学習計画を立てにくいという問題があった。そこで、発表者は、LOD から shallow/deep question を自動生成する技術と、この技術に基づく学習支援システムを開発した。そして、歴史学習の授業にて24 人の被験者を統制群/実験群に分けた実験を行い、システムの有効性を示すことができた。さらに、学習者のシステム利用時間が短く知識の獲得に至らなかった場合においても、学習者が自分の意見を深く考えるようになる傾向がみられるなど、国際的にも通用する成果になっていると判断した。以上より、本論文を研究会優秀賞に推挙する。

11. SIG-LSE:ことば工学研究会

「ニューラルネットワーク物語生成モデル」

浅川 伸一 (東京女子大学)

(SIG-LSE, 第 49 回, 2016年 9 月 25 日)

選考理由:

deep learning を使ってある著者っぽい作品を生成出来る可能性を示した論文である。青空文庫のデータを用いて実験を行っている。deep learning で、ある画家の描いたような絵を描くなどの実験は既になされているが、それを文学作品まで適用したもの。この結果は、新たな作品を生成するなどの、様々なアプリケーションにまで使うことが出来ると考えられ、ことば工学として非常に重要な研究と考え、研究会優秀賞に値すると考える

12. SIG-FIN 金融情報学研究会

「株式掲示板におけるユーザ行動異常検知を用いた相場操縦発見手法に関する研究」

宮崎 邦洋, 松尾 豊 (東京大学工学系研究科技術経営戦略学専攻)

(SIG-FIN-015, 第 15 回 金融情報学研究会, 2015 年 9 月 25 日)

選考理由:

本論文では、株式掲示板において、ユーザの行動異常度を測定することにより、相場操縦行為を発見する手法を提案している。提案手法は、株価の急騰急落を前における掲示板でのユーザの行動を、様々な側面から分析することにより、ユーザの異常行動具合をランキング化するものである。その結果、実際に風説の流布を試みた可能性のあるユーザのリストを作成することに成功し、実務面への応用が可能である点が評価に値すると判断した。研究会での参加者の反応もよく(推薦率:40.6%)、研究会優秀賞にふさわしい論文である。

13. SIG- DOCMAS：データ指向構成マイニングとシミュレーション研究会

「Twitter 位置情報・テキスト情報を用いた人の移動モデル構築と観光地推薦手法の提案」

前田高志ニコラス（東京大学），吉田光男（豊橋技術科学大学），鳥海不二夫（東京大学），大橋弘忠（東京大学）

（SIG-DOCMAS，人工知能学会合同研究会，2015 年 11 月 14 日）

選考理由：

本論文は，tweet に含まれる位置情報を利用して得られた人々の滞在地に関する情報から人の移動モデルを構築し，さらに構築モデルに基づいて地域の魅力を算出する手法を提案したものである。技術的に傑出しているとは評価しにくい，tweet データから現実の人の移動を引き起こす場の魅力をどれだけ読み取れるか挑戦した本研究の試みは，当研究会が求める内容として適しており，日本，イギリス，アメリカと複数の地域を対象として示した実データに基づく結果は十分に興味深いものであった。以上より，本論文は，当研究会の優秀論文候補としてふさわしい内容と認める。

飲食店向け不動産営業を支援する 申込み顧客推薦システムの評価

河村一輝¹, 諏訪博彦¹, 小川祐樹², 荒川豊¹, 安本慶一¹

¹ 奈良先端科学技術大学院大学情報科学研究科

² 立命館大学

kawamura.kazuki.ka3@is.naist.jp

概要 本研究では、飲食店向けの不動産営業を支援する手法を提案する。飲食店起業を目指す顧客が、物件の成約に至る流れは以下の通りである。①検索サイトで不動産を探す。②実際に内見を行い気に入った物件を閲覧する。③内見に行った物件の内気に入った物件を申し込みする。④物件の貸主と交渉し、成約に至る。①→②の検索から内見に移ることや、②→③の内見から申し込みに自然に移ることは稀なため、不動産会社の営業マンが電話をかけて次のステップに促す。従来では電話営業を行う際、営業マンは長年の感と経験によって、数ある顧客リストからより成約に至りやすい顧客を選定し、電話をかけていた。この時の営業成功確率は、2割から3割程度であり、この精度向上は重要な課題である。さらに、新人の営業マンでは、どの顧客に電話をかけるべきかのノウハウがたまっておらず、どの顧客にかけたらいいいのか分からないといった問題があり、その対応策の創出が課題である。この課題を解決するために、本研究では機械学習を用いて、申込みの可能性の高い顧客を推定し、営業マンに電話をかけるよう推薦するシステムを提案する。内見時のアンケート結果、物件情報の基礎データ2クラスの他に熱意と地域ポテンシャルという独自の特徴量を加え、RandomForestを用いて機械学習を行ったところ53.8%の精度で申し込み顧客を推薦することが出来た。ベテランの営業マンの営業成功確率が2割から3割程度のため、本システムで推薦された結果は、営業効率をあげることに大いに貢献出来る。

キーワード 機械学習, 推薦, 営業支援システム

1 背景

近年、様々な実社会の情報を容易に収集出来るようになってきた。また、多くの機械学習に関する研究が行われ、様々な手法が開発されている。それに伴い、それらの情報を使い機械学習を用いて利用者が望む情報を見つける手法[1][2][3]が盛んに提案されている。例えば、Saidehら[4]は、新規の飲食店のオンラインレビューのレイティングを予測しようとしている。

また、これらの手法をビジネスの世界に適用しようとする動きもみられる。具体的には、amazon[5]では、ユーザの商品の購入履歴からユーザが次に購入したがっている商品を推薦し、提案している。また、ビズリーチが提案している転職サイト「キャリアアトレック」[6]では、ユーザの求人情報の閲覧ログからユーザにあった求人情報を提案している。

さらに、機械学習を用いた営業支援研究やシステムも盛んに提案されている。everstring[7]では、B2B 向けの商品を扱っている法人企業を対象にどの法人企業が自社の商品を買ってくれるかを推薦している。また、Yiqing Huangら[8]は、中国の携帯会社のデータを使い、プリペイド携帯の解約者を予測し、その解約者に電話営業を行うよう促しプリペイド解約を思いとどまらせるシステム

を提案している。このように、機械学習を用いて企業の営業を支援し、利益に貢献しようとする多くの取り組みがある。

本研究では、飲食店不動産探索を行う個人に対して営業を行う飲食店不動産会社を対象とし、不動産契約の申込をすると考えられる候補顧客の推薦を行うことで、営業支援を行うことを目的とする。推薦に関する従来研究の多くは、同じ商品を多くの人が購入出来ることを前提としていた。また、比較的安価で、同じ人が複数の商品を購入できることを前提としていた。しかし、本研究が対象とする商品は不動産であるため、世の中に2つとして同じ商品が存在しない。また、複数の不動産を取得するには多額な費用が必要になるため、同一人物が複数回購入することを前提にできない。そのため、従来の推薦手法で多く使われる協調フィルタリングやコンテンツフィルタリングのように、購買履歴から嗜好を生成することや、その嗜好傾向の類似性から商品を推薦することができず、新たな手法を検討する必要がある。

具体的な研究対象としては、飲食店向けの不動産会社である ABC 店舗1と共同研究を行い、データ提供や営業ノウハウについての情報を得ている。ABC 店舗か

1 ABC 店舗 (<http://www.abc-tenpo.com/>)

らの聴き取り調査によれば、飲食店起業を目指す顧客が、物件の成約に至る流れは以下の通りである(図 1)。

- ① 検索サイトにアクセスし不動産物件を探す。
- ② アクセスで気になった物件を内見し、実物件を確認する。
- ③ 内見に行った物件の内、気に入った物件を申し込みする。
- ④ 物件の貸主と交渉し、成約に至る。

しかし、ほとんどの顧客は飲食店を開業することが初めてであり、適切な条件を把握していないケースが多い。こういった顧客が、現行の不動産検索システムを用いて検索を行った場合、適切な条件を入力することが出来ずに、自身に適した不動産を探し成約に至るのは非常に困難である。

そのため、①→②の検索から内見に移る段階や、②→③の内見から申し込みに移る段階で躊躇することがあり、不動産会社の営業マンは電話をかけて次のステップに促している。従来は、電話営業を行う際、営業マンの長年の勘と経験によって、数ある顧客リストからより成約に至りやすい顧客を選定し、電話をかけていた。しかし、新人の営業マンでは、どの顧客に電話をかけるべきかのノウハウがたまっておらず、どの顧客にかけたらいいいのか分からないといった問題がある。また、対応する顧客数の増加に伴い、すべての顧客を手手で検索し続けることは困難である。この問題を解決するために、本研究では機械学習を用いて、申し込みの可能性が高い顧客を推定し、営業マンに電話をかけるよう促すシステムを提案する。本稿では、そのために、どのようなデータを収集し、モデルを構築することで、より効率的なマーケティング支援を行えるのか検討する。

2 飲食店不動産成約プロセス

飲食店不動産を探す顧客の行動プロセスと、各プロセスで得られる情報を図 1 に示す。プロセスは大きく 2 つ、実際に物件を見る前と実際に物件を見た後に分割される。顧客はアクセスステップ、内見ステップ、申し込みステップを経て成約に到る。アクセスステップでは、アクセスログ、お気に入り物件、検索条件といった情報が取得出来る。内見ステップでは、内見ログ、ヒアリング情報といった情報が取得出来る。申し込みステップでは申込ログが取得でき、成約ステップでは成約ログが取得できる。ステップの細かい説明は、以下で述べる。

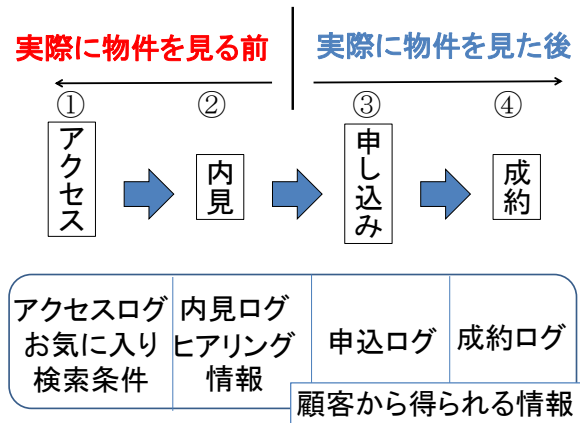


図 1 飲食店不動産契約プロセスと得られる情報

2.1 アクセスステップ

顧客は、検索を行う前に会員登録を行う。その際、会員 ID を付与する。図 2 は、会員登録の情報入力ページの 1 例である。ここでは、氏名、電話番号といった基本的な情報から物件を探す理由(新規出店 or 増店 or 移転 or 顧客への紹介 or その他)、業種(ラーメン or 居酒屋 or イタリアン or 和食 etc)、出店予定時期、勤務状況(勤務中 or 物件探しに専念)、開店資金、開店希望地域、希望する賃料の上限、希望する面積の上限といった物件に対する希望条件を入力してもらい、営業マンは個人の嗜好を把握する。

図 2 登録の物件に対する希望条件入力ページ

その後、ログインを行い図 3 上で物件を検索する。検索方法はキーワード検索である。自身の希望する物件に関するキーワードを入力し検索する。その際、条件を指定するところで更に詳しく検索を行うことが出来る。条件には、地域、最寄りの沿線、業態、賃料の上限、面積の上限、階数、駅からの徒歩分数、飲食条件、居抜き物件かどうかを指定出来る。



図3 検索ページ

図4は検索結果である。このページでは、検索結果にあう条件の物件が表示させる。ここでは、メイン写真、物件番号、最寄り駅、駅からの徒歩分数、階数、面積、賃料(税抜き)、物件に対するキャッチコピー、所在地、業態、保証金といった情報を確認することが出来る。更に、問い合わせボタンをクリックすると物件に関する更に詳しい情報を取得出来るよう営業マンにリクエストを送ることが出来る。また、内見依頼ボタンをクリックすると内見の予約をすることが出来、更にお気に入りボタンをクリックするとABC店舗のwebページ上でその物件に対するお気に入り登録を行うことが出来る。

写真	物件概要	駅	徒歩	階数	面積	賃料(税別)
no image	飲食店で相談可能。入居日即日可能な1階店舗	東京メトロ千代田線 北千住	15分	1階	12.52坪 41.38㎡	100,000円(税別) 押借額：7,989円
	中央通りに面すバー・クラブ層抜き	京山手線 新橋	5分	6階	12.98坪 42.9㎡	363,982円(税別) 押借額：26,048円
	オフィスと住居が融合した街 駅徒歩3分 1階路廊うーめん店層抜き	東京メトロ半蔵門線 水天宮前	3分	1階	9.26坪 30.6㎡	295,000円(税別) 押借額：21,869円
	住宅街と学生街の良いとこ取りエリア！ 駅徒歩4分 1・2階一括貸し居間層層抜き	東武大井町線 大井町	4分	1階~2階一括	129坪 39.66㎡	150,000円(税別) 押借額：12,503円

図4 検索結果ページ

このように、アクセスステップでは、顧客の不動産閲覧履歴であるアクセスログ、気に入った不動産をブックマークするお気に入り登録、不動産を検索する際の絞り込み条件である検索条件をそれぞれ得ることができる。

2.2 内見ステップ

内見ステップは実際に不動産を見るステップである。内見では実際に営業マンが顧客を物件まで連れて行き、顧客が持つ物件に対する印象を会話から読み取って今後の営業につなげていく。また、構造化アンケート形式

で顧客に簡単なヒアリングを行う。

アンケート項目は、「第一印象はいかがですか?」「立地場所はいかがですか?」「譲渡価格はいかがですか?」「取得総額はいかがですか?」「物件がある周辺の場合について事前調査は行いましたか?」である。最初の2項目については、良い、やや良い、普通、やや悪い、悪いの5段階評価で、次の2項目については、高い、やや高い、普通、やや安い、普通の5段階評価で、最後の項目については、行なった、行なっていないの2段階評価である。

このように、内見ステップでは、実際に見た不動産の履歴である内見ログ、内見を行う際に、営業が顧客から直接話を聞いたヒアリング情報を得ることができる。営業は、顧客がその物件を気に入ったかどうか、何が申込みの障害になっているのかなどを聴き取りにより把握している。顧客の中には、繰り返し内見を行う顧客も少なくない。

2.3 申し込みステップ

申し込みは不動産を成約する為の予約をするステップである。申し込みステップでは、顧客が申し込んだ物件履歴である申し込みログを得ることができる。

その際、申し込み情報は貸主に送信される。同時に、他の顧客も同じ物件に申し込みを行うことがあり、その際には複数の申し込みの顧客の中から貸主が物件を貸したい顧客を選定する。そのため、申し込みのステップで、貸主の意向や他の顧客に先をこされるなどの事情で成約に至らなかった場合があり、その顧客はまた内見ステップに戻ることになる。

2.4 成約ステップ

成約は不動産を成約するステップである。申し込みをした不動産貸主と顧客の双方の同意が取れ、手付金を支払うことで漸く成約に至ることのできるものである。成約ステップで、顧客が成約した不動産履歴である成約ログを得ることができる。

2.5 本研究の課題

不動産成約プロセスの中で、内見・申し込み・成約のステップでは、顧客は実際に不動産を見ている状態にあるが、アクセスステップでは不動産をまだ実際には見ていない状態である。不動産を見る前後で顧客のこだわりが変化する為、アクセスステップにある顧客情報のみから、申し込みや成約に至る顧客を予測するのは困難である。また、内見した印象が良かったからといって申し込みをすとも限らない。本研究では、アクセス、内見の両ステップの情報から機械学習を用いて申し込み顧客を推定することを試みる。次章において、機械学習による人間の行動

予測に関する研究を紹介する。

3 関連研究

機械学習を用いた商品推薦,営業支援,不動産の価値推定に関してこれまで様々な研究が行われている。本研究の位置づけを示すために,関連研究を以下に説明する。

3.1 商品推薦の研究

Amazon[5]では, EC サイトで商品を購入する顧客に対してユーザの購入ログから, よりユーザが好む商品の予測を行っている。Amazon[5]を利用するユーザは, 商品の購入後,5 段階(好き,少し好き,普通,少し嫌い,嫌い)の評価を行う。システムは, その評価を元に協調フィルタリングによってユーザ間の類似度を求める。ユーザが高い評価を行った商品は, 類似度の高い他のユーザにも推薦される。

ビズリーチの転職サイト「キャリアトレック」[6]では, 会員の基本情報や診断テスト, サイトの利用動向, 似ているユーザのサイトの利用状況からそのユーザに合った求人情報を提案している。このシステムでは, コンテンツベースフィルタリングを用いている。まず, システムは会員登録時基本情報からユーザの嗜好を表すプロフィールベクトルを作成します。その後, ユーザのプロファイルベクトルと会社の求人情報が有するプロフィールベクトルとの類似度を計算し転職すべき会社を推薦する。

3.2 機械学習を用いた営業支援の研究

everstring[7]では,B2B の企業に対して,商品を買ってくれそうな顧客を推薦するサービスを行っている。従業員規模,収益状況,経営者の経歴,地域などを説明変数に用いて,成約確度の高いもしくは低い顧客を判定するモデルを機械学習で構築している。YiqingHuang ら[8]の研究では,中国の携帯会社の営業支援システムを提案している。プリペイド携帯の解約者を予測し,営業マンに電話をかけさせることで解約を思い留まらせる仕組みである。ユーザの行動ログ,メッセージログ,パケット,通話の品質等から解約する確率の高いユーザを推定している。

3.3 不動産営業支援に関する研究

Yanjie ら[10]の研究では, dianping[11]のレーティングやチェックイン情報等のオンラインの情報とタクシーやバスのトラッキングなどのオフライン情報から機械学習を行い,中国の不動産価格の推定を行なっている。チェックインの回数や物件ごとのレーティングが高い場所が不動産の価格推定において大きな重要度を示すことが分かった。chin ら[12]の研究では,物件の面積や築年数といった物件情報だけでなく風水の属性を入れたデータセットを用いて不動産の価格推定を行なっている。全体を通し

ての価格推定の重要度は物件の延べ床面積が一番高いという結果になった。また,風水の属性を含めたデータセットは,風水の属性を除いたデータセットと比べてより高い精度を出すことが出来,風水の属性の重要性を示している。本手法では,物件の良し悪しが物件の申込み行動に影響を及ぼすと考え,上記関連研究の属性を参考にデータセットを作成する。得田ら[13]や田邊ら[14]の研究によると不動産の価格は物件の良し悪しに依存しているため,上記の研究は本研究に適用出来る。安藤ら[9]の研究では,ABC 店舗において1→2の検索から内見に移るステップを支援するために,顧客のこだわりを抽出する手法を提案している。この手法では,顧客のアクセスログ,お気に入り,検索条件から,顧客が7つの属性(場所,駅徒歩,階数,面積,賃料,居抜き,造作)のうちどの属性にこだわりがあるかを算出し,それにより顧客のプロフィール(こだわり)を作成し,こだわりと物件のマッチ度が高い顧客ほど内見しやすいと考えて推薦している。しかし,実際に申し込みをするかどうかについては検討されていない。そこで,本手法では2→3の内見から申し込みに移るステップを支援する手法を提案する。

3.4 本研究の位置づけ

推薦に関する従来研究の多くは,同じ商品を多くの人が購入出来ることを前提としていた。また,比較的安価で,同じ人が複数の商品を購入できることを前提としていた。しかし,本研究が対象とする商品は不動産であるため,世の中に2つとして同じ商品が存在しない。また,複数の不動産を取得するには多額な費用が必要になるため,同一人物が複数回購入することを前提にできない。そのため,従来の推薦手法で多く使われる協調フィルタリングやコンテンツベースフィルタリングのように,購買履歴から嗜好を生成することや,その嗜好傾向の類似性から商品を推薦することができず,新たな手法を検討する必要がある。

4 提案手法

図5のように従来手法において,営業は手作業で大量の内見済みの顧客の中から不動産の申込みの可能性が高い顧客の発掘を行っている。そのため,手間と労力が必要であり,コストがかかっている。また,営業の成功確率は,営業の勘と経験から導き出されるノウハウ(暗黙知)により変化する。そのため,ノウハウ(暗黙知)を形成できていない新人営業は,営業機会を損失していることが考えられる。そこで,提案手法では暗黙知であるノウハウを機械学習によりモデル化することを目的とする。また,その際に,どのようなデータがモデル構築に必要なのか検証する。

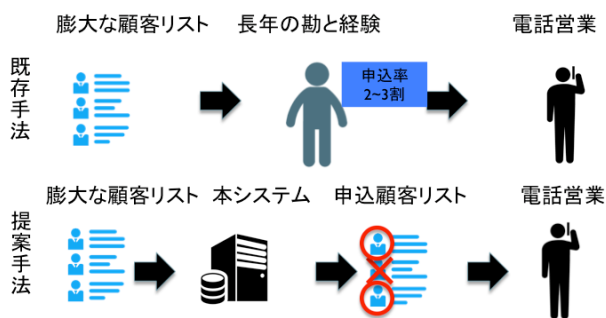


図5 既存手法と提案手法の違い

モデル構築のためのデータセットとして、大きく3つのクラス(属性群)を用いた。1つ目のクラスは、内見後に顧客に答えてもらったアンケートデータである。アンケートでは、第一印象、立地、月額賃料、譲渡価格、取得総額について顧客がどう感じているかと事前調査の有無について確認した。顧客が申し込みをするか否かは、顧客が持っている物件に対する主観評価が申し込み行動に影響を及ぼすと考え、このクラスをデータに含めることとしている。

2つ目のクラスは、物件の情報である。物件の情報には、月額賃料、面積、何階か、駅から徒歩何分か、最寄り駅、居抜き物件かどうかが含まれている。ABC店舗が扱う物件は多種多様であり、物件立地や価格などの物件自体の特長により、人気・不人気が存在する。このことから、物件自体の特徴が顧客の申し込み行動に影響を及ぼすと考え、このクラスをデータに含めることとしている。

3つ目のクラスは、基礎データではなく加工した特徴量群である。開店意欲(量)、開店意欲(期間)、地域ポテンシャルが含まれている。開店意欲をもって真剣に物件を探している顧客は、物件を申し込みの可能性が高いというABC店舗の知見により、この特徴量を作成する。開店意欲には、2種類があると考えている。まずは、意欲の大きさである。これは、どれだけ大きな意欲を持っているかという絶対的な量を表す。より意欲の高い顧客は様々な物件にアクセスするという仮定のもと、ABC店舗のデータベースからアクセスログを取得し内見日の14日前から内見日までの2週間での総アクセス数を開店意欲(量)と定義した。

次に、意欲の継続性である。これは、どれだけ長い時間開店意欲を継続できているかを表す。より長く継続して物件を探索している人ほどより意欲が高いという仮定のもと、ABC店舗のデータベースからアクセスログを取得し内見日の14日前から内見日までの総アクセス日数を開店意欲(期間)と定義した。

地域ポテンシャルについては、物件がある街の人気度は、申し込み行動に影響を及ぼすと考え、街全体の平均賃料が高いほど地域ポテンシャルも高いと仮定し、データベースから物件情報のデータを取得し最寄り

駅毎の単位面積当たりの平均賃料を地域ポテンシャルと定義した。

提案システムでは、これら全てのデータを組み合わせることで申し込み推定モデルを構築し、申し込み見込み顧客リストを構築する。

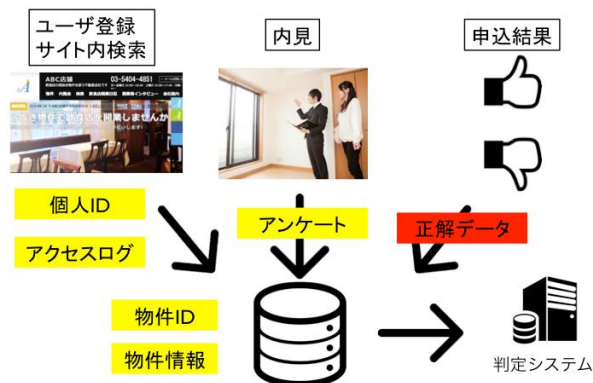


図6 データの取得とモデル

5 システム構成

本システムは、図7のように説明変数のデータをまとめるデータ処理機構、ABC店舗の個人情報、物件情報、アクセスログをテーブルに持つデータベース、1132件のトレーニングデータからRandomForest [15]によりモデルを作った申し込み顧客推定モデルから構成される。

1. トレーニングデータから申し込み顧客分類モデルを構築する。
2. その後、電話営業マンは顧客リスト(個人ID、物件ID、アンケート結果、内見日が記入されたもの)をcsv形式で読み込ませる。
3. 入力された個人ID、物件ID、内見日から説明変数に必要なデータを取得するためデータベースにクエリを送信する。
4. クエリの結果(物件データ、地域ポテンシャル、熱意)をデータ処理機構に送信する。
5. データ処理機構でまとめた説明変数を申し込み顧客モデルに送信する。
6. 2値(申し込みをする顧客、申し込みをしない顧客)分類を行う。
7. 個人IDから電話番号を取得するクエリを送信する。
8. 電話番号を取得し、個人ID、物件ID、電話番号、予測結果のリストを表示する。

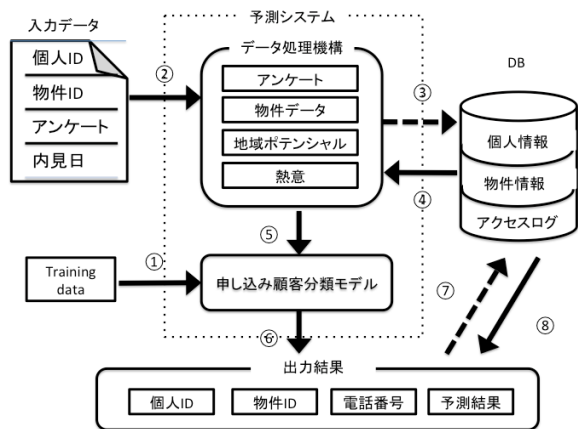


図7 申し込み顧客推薦システムの構成

6 実験

提案手法で示したデータから申し込み推定モデルについて検証するために、実際にデータを収集して評価実験を行った。本研究で使用するデータセットを表1に示す。アンケートデータは、不動産会社の営業の方にとって頂き、顧客が物件を内見した後に構造化インタビュー形式でデータを収集している。2014年2月3日から7月10日まで、計1146件のデータを収集している。

これらのデータセット(表1)を用いて、機械学習により申し込み推定モデルを構築し、交差検証法を用いて推定結果の精度を検証した。モデル構築には代表的な機械学習アルゴリズムの一つである RandomForest [17]を用いた。モデル構築に際して、様々な条件下でデータを組み合わせ、計5つのモデルを構築している。5つのモデルは、1. アンケートデータクラスのみを用いたモデル、2. 物件情報クラスのみを用いたモデル、3. 基礎データ(アンケートデータクラスと物件情報クラスを合わせたもの)を用いたモデル、4. 表1の全てのデータを用いたモデル、5. より重要なファクターのみを抽出した場合モデルである。重要なファクターの抽出は、ジニ係数の平均減少量(mean decrease gini)を用いて行なっている。ジニ係数は下記式で求めた。

$$Gini = \frac{\frac{1}{2} - \int_0^1 L(F) dF}{\frac{1}{2}}$$

表1 本研究で使用するデータセット

クラス	属性
アンケート	第一印象
	立地
	月額賃料
	譲渡価格
	取得総額
	事前調査
物件情報	月額賃料
	床面積
	駅徒歩
	階数
	居抜き
	お気に入り
加工した特徴量	開店意欲(量)
	開店意欲(期間)
	地域ポテンシャル

7 実験結果

本章では、各データセットを用いた際の推定結果について述べる。まずは初期実験として簡単に使用出来る基礎データを用いて検証をした。表2は、アンケート結果のみを用いて申し込み顧客を推定した結果である。申し込みをすると推定した人数98人中、実際に申し込みをした人数は、30人であり、30.6%の精度で推定できた。

表3は、物件情報のみを用いて申し込み顧客を推定した結果である。申し込みをすると推定した人数59人中、実際に申し込みをした人数12人であり、20.3%の精度で推定できた。アンケート情報、物件情報で比べた場合、アンケート情報を用いた時に精度が一番高い結果が出た。物件の申し込みには、心理状態が他の要素と比べて大きく影響していることが分かる。

表4は、物件情報とアンケートを用いて申し込み顧客を推定した結果である。申し込みをすると推定した人数58人中、実際に申し込みをした人数22人であり、37.9%の精度で推定できた。物件情報のみを用いた場合が20.3%の精度であり、アンケートのみを用いた場合では30.6%であったため、7.3ポイント精度を向上することが出来た。

以上の結果、アンケートデータと物件データの両方を用いた際に、一番良い結果が出る事が確認できた。次に精度を改善するために、ABC店舗の知見に基づいて加工した特徴量である開店意欲(量)、開店意欲(期間)、地域ポテンシャルの3つの属性を加えて推定を実施した。表5は表1の全データを用いて申し込み顧客を推定

した結果である。申し込みをするだろうと推定した人数 50 人中、実際に申し込みをした人数 22 人であり、44.0%の精度で推定できた。全ての基礎データを用いて申し込み顧客を推定した結果が 37.9%であったため新たな特徴量を作成することにより 6.1%精度を向上することが出来た。

更なる改善のために、表1の特徴量の内影響力のある特徴量を調べるためジニ係数の平均減少量を計算した。その結果が図 8 である。図を見て分かるように新たに作成した特徴量、地域ポテンシャル、開店意欲(量)、開店意欲(期間)の重要度はそれぞれ1番、6番、8番となった。よって、基礎データだけでなくABC店舗の知見を元に新しい特徴量を作成することは有用であった。この結果を用いて平均ジニ係数の減少量が低い特徴量から1つ1つ要素を削っていき特徴量の数がいくつの時に最大の成果が得られるのかを検証した。その結果6つの要素を削った時、即ち上位9個の特徴量を用いた際に、予測結果が一番高かったため今回は上位9個の変数を使い求めた。

その結果が表6である。申し込みをするだろうと推定した人数 52 人中、実際に申し込みをした人数 28 人であり、53.8%の精度で推定できた。初期実験で行った心理状態を表すアンケート結果のみの場合と比べて、最終的に 23.2 ポイント精度を向上することができた。

表2 アンケートから推定した結果

	申込まない	申込む	
申込してない	889	68	93.0%
申込みした	136	30	18.1%
精度	86.9%	30.6%	

表3 物件情報から推定した結果

	申込まない	申込む	
申込してない	920	47	95.1%
申込みした	153	12	7.2%
精度	85.7%	20.3%	

表4 アンケート、物件情報から推定した結果

	申込まない	申込む	
申込してない	931	36	96.2%
申込みした	143	22	13.3%
精度	86.7%	37.9%	

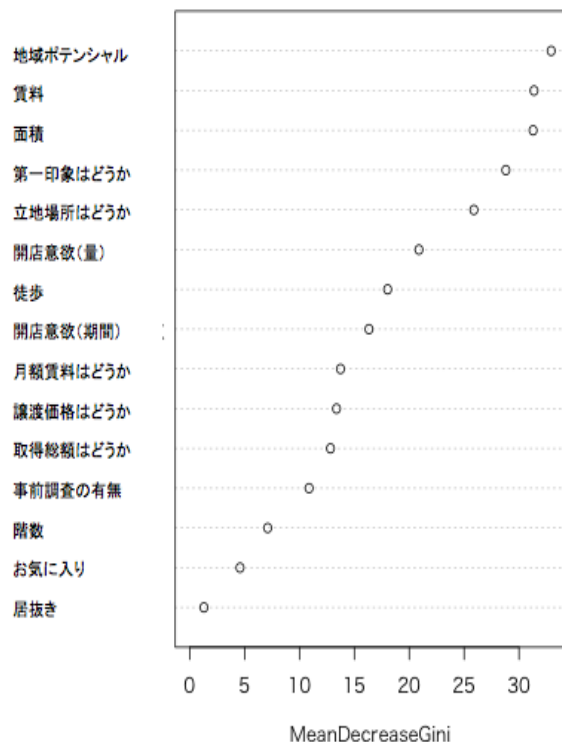


図 8 各特徴量の平均ジニ係数減少量

表5 表1の全データから推定した結果

	申込まない	申込む	
申込してない	939	28	97.1%
申込みした	143	22	13.3%
精度	86.8%	44.0%	

表6 図8の上位9個の特徴量から推定した結果

	申込まない	申込む	
申込してない	943	24	97.5%
申込みした	137	28	17.0%
精度	87.3%	53.8%	

8 考察

5つのシナリオに対する推定結果について、データ提供企業であるABC店舗からのフィードバックも踏まえて考察を行う。なお、ABC店舗からのフィードバックは、推定結果の提示・説明および質疑応答の会議形式で取得している。

5つのシナリオを比較した結果、全体としては全てのクラスデータの上位9個のデータを使用した際に最も効率よく申し込み顧客を推定できることが分かった。推定精度は53.8%であり、理論上は、申し込み候補者リストのうち約半分が申し込みをすることとなる。この結果をABC店舗に伝えたと、従来の営業手法ではベテランの営業マンでも2~3割程度の精度であり、本手法は営業の

効率化に寄与すると思うとのコメントを得ている。

また、地域ポテンシャルについては、「良い立地の物件は申し込みが殺到する。良い物件は顧客の状況に関わらず申し込みされる」とのコメントが得られており、我々の分析結果と対応していると考えられる。

さらに、内見時の顧客の反応としてどのようなところに注しているかを確認した所、顧客の言葉として「立地」「賃料」という言葉が出てきた時に申し込みの有無が判断しやすいとのコメントが得られている。分析結果では、賃料が2番目、立地が5番目に重要なファクターとして抽出されており、この結果についても整合性が取れていると考える。

今後の課題としては、実営業の中で評価実験を行うことでその効果を定量的に図る必要がある。具体的には、提案手法を用いて抽出した顧客を含めた内見顧客に対して、実際に電話営業を行いその反応結果を取得し、比較する必要がある。その際には、既に申し込みをしている顧客の回避やダブルブッキングの回避などのフィルタリングを行うなど、綿密な実験計画が必要となる。

9 まとめ

申し込み見込顧客を営業担当者に推薦することで、営業担当者の時間とコストを省く推薦手法を提案した。申し込みをしたいと思う顧客の行動を予測するためには、地域ポテンシャルや物件の賃料、面積など人間の主観的な気持ちというよりも商品そのものの価値が大きな影響を与えていることが分かった。また、データを取得出来る種類が少ない際には、基礎データを加工し新しい特徴量(地域ポテンシャル、開店意欲等)を作成することは、有用であった。初期実験のアンケート結果を用いて機械学習を行った精度は、30.6%であったが、全てのデータを用いて機械学習すると53.8%の精度で申し込み顧客を推薦することが出来た。

今後はさらなる精度向上を図るために、顧客が内見に来訪している状態や営業マンの身体データセンシングすることで、申し込み顧客の推定ができないか検討する。また、本システムを使用し実環境で評価実験を行う際、営業マンが使いやすいシステムでないとかえって手間が増えてしまう可能性があるため、使いやすいユーザーインターフェースを検討し改良する。今回は、顧客が物件を申し込みか申し込まないかの2値を予測したが、ABC店舗から2値ではなく申し込み見込確率を知りたいとの要望があったため見込確率が高い順にソートされた顧客リストが出力できるシステムを作成する。

謝辞

株式会社ABC店舗の方には、本研究において、データ提供並びに貴重なご意見を頂きました。心より御礼

申し上げます。本研究の一部は科学研究費補助金 挑戦的萌芽研究 15K12161 の助成を受けたものである

参考文献

- [1] 神嶋 敏弘,推薦システム—情報過多時代をのりきる,情報の科学と技術,56巻10号,pp.452-457,2006
- [2] J.Ben Schafer,Joseph A.Konstan and John Riedl,“E-Commerce Recommendation Applications,” *Data Mining and Knowledge Discovery*, Vol. 5, pp.115-153, 2001
- [3] 土方嘉徳, 情報推薦・情報フィルタリングのためのユーザープロファイリング技術, 人工知能学会誌, Vol.19, No.3, pp.365-372, 2004
- [4] Saeideh Bakhsi, Partha Kanuparth, Eric Gilbert “Demographics, Weather and Online Reviews: A Study of Restaurant Recommendations” *Proceedings of the 23rd international conference on World wide web* p.p 443-454
- [5] amazon <<http://www.amazon.co.jp/>>
- [6] キャリアトレック <<https://www.careertrek.com/>>
- [7] everstring <<http://www.everstring.com/>>
- [8] Yiqing Huang Fangzhou Zhu Mingxuan Yuan Ke Deng Yanhua Li Bing Ni Wenyuan Dai Qiang Yang and Jia Zeng, “Telco Churn Prediction with Big Data”, *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp 607-618,
- [9] 安藤孝, 小川祐樹, 諏訪博彦, 太田敏澄, 飲食店不動産の営業を支援する内見顧客推薦システムの提案と評価, 第19回社会情報システム学シンポジウム学術講演論文集, pp.129-135, 2013.
- [10] Y. Fu, Y. Ge, Y. Zheng, Z. Yao, Y. Liu, H. Xiong, and N. J. Yuan, “Sparse real estate ranking with online user reviews and offline moving behaviors,” in *Data Mining (ICDM), 2014 IEEE International Conference on*. IEEE, 2014, pp. 120–129.
- [11] “Dianping,” <<http://www.dianping.com/>>
- [12] C.-H. Wu, C.-H. Li, I.-C. Fang, C.-C. Hsu, W.-T. Lin, and C.-H. Wu, “Hybrid genetic-based support vector regression with feng shui theory for appraising real estate price,” in *Intelligent Information and Database Systems, 2009. ACIIDS 2009. First Asian Conference on*. IEEE, 2009, pp. 295–300.
- [13] 得田雅章, “不動産価格と実体経済: 住宅地地価に関するファンダメンタルズ・モデルの妥当性,” 2014.
- [14] 田邊信之, “第176回定期講演会 講演録不動産関係者のための最新情報(教養講座),” *土地総合研究*, vol. 22, no. 2, pp. 140-182, 2014.
- [15] Leo Breiman, “Random Forests”, *Machine Learning October 2001*, Volume 45, Issue 1, pp5-32, 2001

実データに基づく宅配便物流の地域集約方式による効率化の研究

Simulation Study with Real Data to Improve the Performance of Home Delivery Service Industry

黒瀬雄太^{1*} 寺野隆雄¹
Kurose Yuta¹ Terano Takao¹

¹ 東京工業大学 大学院総合理工学研究科 知能システム科学専攻
¹ Department of Computational Intelligence and Systems Science,
Tokyo Institute of Technology

Abstract: In logistics industry, long-distance transportation at night is currently employed. But, to deal with the increase in the number of buggages, they need to introduce the operation around the clock in order to shorten the delivery time. This paper quantitatively investigates what are the possible issues and the solutions for them by simulation approach. The main findings are as follows: First, the delivery around the clock is much more costly than that at night. Second, by introducing "gateway" among several bases, which can consolidate the buggages in local at first and then transport them to the other regions, shorter delivery time and lower delivery cost have been observed.

1 序論

電子商取引 (Electronic Commerce : EC) 市場の拡大に伴い, 日本国内における宅配便取扱量は年々増加している [1][2]. この EC 市場関係の荷物は「お急ぎ便」や「当日配送」といった時間的制約を持つものが多く, また注文した商品を早く受け取りたいという顧客のニーズもますます高まっている. この時間的制約は宅配便事業者の信頼度に直接関わる問題であるため, 各事業者は更なる輸送時間短縮の必要性に迫られている [3].

この問題に対応するため, 宅配便事業者であるヤマト運輸株式会社 (以下, ヤマト) では, 超大型物流施設を介して輸送時間を短縮する「ゲートウェイ (GW) 構想」を掲げ, 現状では達成していない地域を越えた当日輸送の実現による輸送時間の短縮を試みているが, GW 構想の実現に際し, 具体的な金額コストや輸送時間などの定量的な評価は行われていない [4][5].

これまでに輸送拠点内の効率的なレイアウトや宅配便配送ルート最適化, パレットの運用を考慮した輸送方法など, 宅配便の幹線輸送以外の部分に着目した効率化に関する研究は行われている [6][7][8]. また, 関連研究として, 輸送効率の向上, コスト削減といった一般的な研究は数多く行われているが, 本論文のよう

な特定のスケジューリング問題に対しては適応が難しい [9][10][11].

本論文では実データに基づくシミュレーションにより, はじめに現状の夜間みの幹線輸送を単純に 24 時間化した場合に発生する問題点を定量的に評価し, 次に対処法として考えられる地域集約型 (GW) 方式の幹線輸送を用いた場合の金額コストや輸送時間などといった評価値の変化を検証することを目的とする.

2 シミュレーション

本章では輸送方式とシミュレータの説明やパラメータなどの各設定を行う. なお, 説明で用いられる「BOX」はロールボックスパレットの略称であり, 輸送拠点間で輸送される荷物の最小単位である. また, シミュレーションにおける 1 ステップ (step) を実時間の 30 分とする. なお, 高速道路の夜間割引や, 渋滞による所要時間に差をつけるため, 「日中」と「夜間」を定義し, 1 ステップ目を夜間の開始時間, 25 ステップ目を日中の開始とし, 以降 24 ステップごとに日中と夜間を繰り返す.

2.1 輸送方式

シミュレータの説明を行うにあたり, 図 1 を用いて現在の輸送方式を説明する. ヤマトでは図 1 のような 3 層からなる多段階輸送方式を採用しており, 送り主か

*連絡先: 東京工業大学 大学院総合理工学研究科 知能システム科学専攻

〒 226-8502 神奈川県横浜市緑区長津田町 4259 J2-1705

E-mail: kurose.y.aa@m.titech.ac.jp

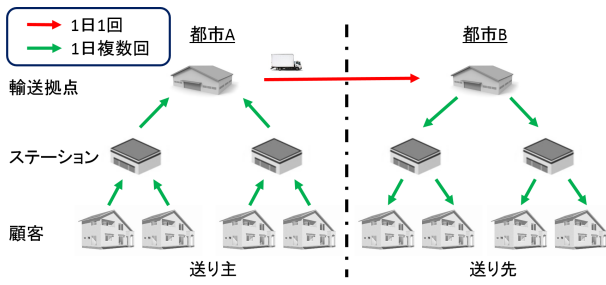


図 1: 現在の輸送方式

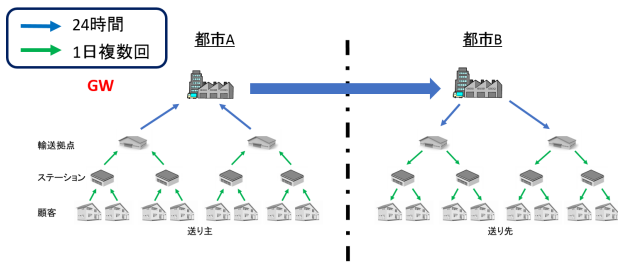


図 2: GW 構想での輸送方式

らの荷物はステーション、輸送拠点を通して送り先へと輸送される。各層を結ぶ輸送はそれぞれトラックの大きさや輸送頻度が異なり、顧客ステーション間およびステーション輸送拠点間は1日に複数回輸送が行われる。しかしながら、輸送拠点間の輸送にあたる幹線輸送は1日に1度夜間のみしか行われておらず、これにより現状では地域を越えた当日輸送は実現していない。

次に、図2を用いてGW構想での輸送方式を示す。GW構想での輸送方式は図1で示した3層の輸送方式のさらに上層に、GWという超大型物流施設を配置した4層からなり、輸送拠点間とGWおよびGW間は24時間輸送が行われる。これにより地域を越えた当日輸送が実現可能となる。

今回のシミュレーションでは、これらの輸送方式の変更に伴う金額コストや輸送時間を比較するため、対象となる輸送拠点から上の層のみに着目する。なお、図1で示した幹線輸送が夜間にのみ行われるモデルを「夜間モデル」、夜間モデルを単純に24時間化したモデルを「24時間モデル」、図2で示したGW構想での輸送方式を「GWモデル」とする。

2.2 シミュレーションの概要

出発地側の輸送拠点にBOXが発生してから目的地側の輸送拠点に到着するまでの幹線輸送のみに着目し、輸送拠点間の輸送の実データから平均輸送時間や金額コストといった評価値を導くモデルを用いたシミュレ

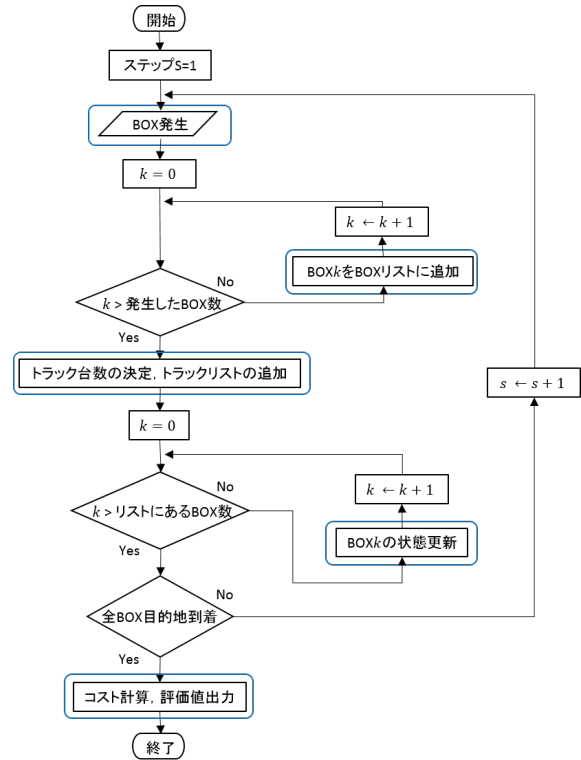


図 3: シミュレーションのフローチャート

ーションを実装する。内部では行列計算によりBOXを移動させ結果を得る。またBOXの輸送可能条件や経由方法をモデルごとに変更し実験を行う。

シミュレーションには輸送拠点をノードとした輸送拠点間の物流ネットワークを用いる。輸送拠点は全国に69カ所存在し、各輸送拠点間の距離および所要ステップ数はあらかじめ与えられている。

なお、輸送拠点間の距離および所要ステップ数は、実際の地図情報から配車計画用ソフト「Lyna2」[12]を用いて算出したデータを使用する。

2.3 フローチャート

本節では図3を用いてシミュレーションの主な流れを示す。

はじめに発側の輸送拠点にBOXが発生すると、BOXごとにBOXリストへ追加する。このBOXリストに「BOXの状態」を記録し、これを更新することでBOXの管理を行う。発生するBOXは実データに基づきステップごとに割り振ったものであり、出発地、目的地と発生時間が与えられている。

次に、BOXリストから各方面へのBOX数を算出し、その値から、トラックの台数を算出する。今回のシミュレーションでは、積載率が100%となればトラッ

クが発車可能となる。輸送拠点 i から輸送拠点 j へのトラック台数 T_{ij} [台] は、 i から j へ輸送予定の BOX 数を n_{ij} [BOX]、トラック 1 台あたりの最大積載 BOX 数を L [BOX] とすると次の式で表される。

$$T_{ij} = \lfloor n_{ij}/L \rfloor \quad (1)$$

これによりトラック台数が決定すると、各トラックをトラックリストへ追加する。トラックリストへは、発車するトラックごとに出発地、目的地、移動距離 d_{ij} [km]、稼働時間 t_{ij} [step] や積載率を記録する。次にトラック台数の決定に伴い、輸送 BOX を決定する。トラックへの積載可能 BOX 数 α_{ij} [BOX] は、式 (1) を用いて次の式で与えられる。

$$\alpha_{ij} = L \times T_{ij} \quad (2)$$

また式 (1) および式 (2) より、待機となる荷物数 W_{ij} [BOX] は次の式で与えられる。

$$W_{ij} = n'_{ij} - \alpha_{ij} \quad (3)$$

これらの値に応じて BOX ごとに、発車、待機を決定する。このとき、トラックへの積載は、前ステップで待機していた BOX から優先的に行われる。

発車となった BOX はその BOX の「現在の状態」を「移動中」に更新し、移動に必要なステップ数を与える。このステップ数を消化すると、目的地へ到着となる。なお、発車となる条件には上記の他に夜間モデルでの日中の待機や強制発車がある。強制発車は発車できずに待機となったステップ数が、最大待機ステップ数 β [step] となった場合に強制的にトラックを発車させるものであり、この場合、当該トラックを新たにトラックリストへ追加する。またこの場合、同一方面へ待機中の BOX (W_{ij} [BOX]) 全てが積載される。

ここまでの操作をステップごとに行い、発生した全 BOX の輸送が完了すると BOX リストおよびトラックリストから各評価値を出力すると共に金額コストの計算を行う。金額コストは、トラックごとにかかる人件費 C_P [円]、高速代 C_H [円] および燃料代 C_F [円] のみ考慮し、これ以外の固定費などの費用は考慮しない。なお、各費用は次の式で定義される。

$$C_P = t_{ij} \times w/2 \quad (4)$$

$$C_H = d_{ij} \times h \times r \quad (5)$$

$$C_F = d_{ij} \times 1/f \times g \quad (6)$$

ここで、 w [円/h] はドライバーの平均時給を表し、2 で割ってステップ単位とする。また、 h [円/km] は高速料金であり、 r は夜間のステップのみ適用される深夜高速料金割引である。また、 f [km/L] はトラックの燃費であり、 g [円/L] は軽油代を表す。式 (4)、式 (5)、式

(6) より、1 台ごとにかかる金額コスト C [円] は次の式で与えられる。

$$C = C_P + C_H + C_F \quad (7)$$

次に、出力する評価値を説明する。

- BOX 数 [BOX]
総 BOX 数を表す。
- トラック台数 [台]
輸送に要したトラック台数を表す。なお、同一トラックの複数回使用を考慮していないため、全て「のべ台数」である。
- 平均輸送時間 [step/BOX]
出発地側の輸送拠点に BOX が発生してから、目的地側の輸送拠点に BOX が到着するまでにかかる BOX あたりの輸送時間を表す。
- 平均積載率 [%/台]
トラック 1 台あたりの平均積載率を表す。
- トラック平均移動距離 [km/台]
トラックごとの平均の移動距離を表す。
- トラック平均稼働時間 [step/台]
トラックが出発地から目的地に到着するまでの所要時間のトラックあたりの平均を表す。トラックの移動時間のみを計測しているため、待機時間は含まれない。
- BOX キロ (合計)[百万 BOXkm]
物流業界でトラックの輸送コストの指標として広く用いられる「トンキロ」を基にしたコストである。トンとキロの積で与えられるトンキロと同様、トラックごとの、積載 BOX 数と移動距離の積で表され、全トラックの BOX キロの合計を表す。
- 金額 (合計)[百万円]
トラックごとに算出した輸送にかかる金額コストの合計を表す。

以上がシミュレーションの流れである。

2.4 条件設定

- GW
GW モデルでは、全輸送拠点をヤマトの地域に従い 9 地域に分割し、各地域に必ず 1 つ GW を配置した。各地域の GW の決定方法は、GW 同士の距離を短くするため、既存の輸送拠点のうち、他の全拠点との距離の総和が地域内で最も小さい輸送拠点を GW とした。具体的に例を示す

表 1: GW 管轄表

地域	GW 所在地	輸送拠点数	BOX 数
北海道	北海道	5	7225
東北	福島	6	14062
関東	神奈川	18	90049
北信越	長野	7	14936
中部	三重	7	23150
関西	滋賀	9	42532
中国	岡山	5	10151
四国	徳島	4	5249
九州	福岡	8	17076

表 2: 各文字の値

文字	定義	値
β	待機可能最大ステップ数	6
w	ドライバー平均時給 [円/h]	1804[13]
h	高速料金 [円/km]	40.6[14][15]
d	夜間割引 [倍]	0.7[16]
f_{10}	10t 車燃費 [km/L]	6.00[17]
L_{10}	10t 車最大積載 BOX 数 [BOX]	16
f_{20}	20t 車燃費 [km/L]	4.04[17]
L_{20}	20t 車最大積載 BOX 数 [BOX]	24
g	ガソリン代 [円/L]	72.7[18]

と、九州地方であれば福岡県のように各地域の門 (Gate) となる位置に GW を配置した。各地域の GW の配置及び管轄輸送拠点数、取扱い BOX 数を表 1 に示す。なお、同一地域内への輸送は直接目的地へと送られ、地域を越えた輸送には必ず出発地側、目的地側の GW をそれぞれ経由するものとする。

- **トラック**
GW モデルの GW 間の輸送のみ 20t トラックを使用し、それ以外の輸送はすべて 10t トラックを使用する。10t トラックの最大積載 BOX 数を L_{10} [BOX]、燃費を f_{10} [km/L] とし、20t トラックの最大積載 BOX 数を L_{20} [BOX]、燃費を f_{20} [km/L] とする。
- **パラメータ設定**
シミュレーションを行う上で必要なパラメータの定義と設定を行う。具体的な値は表 2 のとおりに設定した。なお、 w と g は経済産業省の統計値を用い、 h は国土交通省と全日本トラック協会の値から算出した。また、 r は NEXCO 中日本、 f は経済産業省より引用した。さらに、今回の実験では、日中の渋滞を表現するため、日中の時間表を夜間の 1.07 倍とした。なお、この値は国土交通省の統計値 [19] より、日中の混雑時と非混雑時の差から算出した。

表 3: 実データと夜間モデルの比較結果

	実データ	夜間モデル
BOX 数 [BOX]	164337	164337
平均輸送時間 [step/BOX]	-	34.4
トラック台数 (計)[台]	11197	11243
平均積載率 [%/台]	91.7	91.4
トラック平均移動距離 [km/台]	285.5	311.8
トラック平均稼働時間 [step/台]	10.2	9.7
BOXkm(合計)[百万 BOXkm]	49.7	49.7
金額 (合計)[百万円]	233.0	240.2

3 予備実験

3.1 シミュレータの精度確認

はじめに、実データから算出した各値と、現状と同じ幹線輸送方式にあたる夜間モデルのシミュレーション結果を比較し、作成したシミュレータの妥当性を検証した。実データと夜間モデルの比較結果を表 3 に示す。なお、実データの金額コストは実際の輸送データのトラックごとの稼働時間および輸送拠点間の距離データをもとに、2.3 で示した金額コストの式を用いた算出した。この比較結果から、金額の誤差が 3%程度であるため、シミュレータの精度は十分であると考えられる。

3.2 夜間モデルと 24 時間モデルの比較

現状の幹線輸送方式を単純に 24 時間化した場合に発生する問題を定量的に評価するために、夜間モデルと 24 時間モデルの比較を行った。比較結果を表 4 に示す。夜間モデルでは発側の輸送拠点で夜間まで待機していた荷物が、24 時間モデルではすぐに発送されるようになったことにより、平均輸送時間が夜間モデルの半分以下に短縮された。しかしながら、トラックが積載率の低い状態で発車することになり、トラック台数が増加したため、金額コストが大幅に増加する結果となった。さらに、24 時間モデルではトラックの平均移動距離が 100 キロ近く長くなっており、ドライバーの負担も増大していることから、単純に現在のモデルを 24 時間化しただけでは当日輸送の実現は難しいことが判明した。

4 実験

4.1 24 時間モデルと GW モデルの比較

24 時間モデルと GW モデルの比較を行い、24 時間モデルで生じた問題点と各評価値の変化を検証した。比較結果を表 5 に示す。

表 4: 夜間モデルと 24 時間モデルの比較結果

	夜間モデル	24 時間モデル
BOX 数 [BOX]	164337	164337
平均輸送時間 [step/BOX]	34.4	12.0
トラック台数 (計)[台]	11243	19414
平均積載率 [%/台]	91.4	52.9
トラック平均移動距離 [km/台]	311.8	402.2
トラック平均稼働時間 [step/台]	9.7	12.5
BOXkm(合計)[百万 BOXkm]	49.7	49.7
金額 (合計)[百万円]	240.2	535.1

表 5: 24 時間輸送モデルと GW モデルの比較結果

	24 時間モデル	GW モデル
BOX 数 [BOX]	164337	164337
平均輸送時間 [step/BOX]	12.0	14.1
トラック台数 (計)[台]	19414	18602
トラック台数 (10t)[台]	19414	14964
トラック台数 (20t)[台]	0	3638
平均積載率 [%/台]	52.9	94.4
トラック平均移動距離 [km/台]	402.2	177.2
トラック平均稼働時間 [step/台]	12.5	5.6
BOXkm(合計)[百万 BOXkm]	49.7	61.0
金額 (合計)[百万円]	535.1	237.5

4.2 考察

表 5 のとおり、GW モデルでは、GW を経由するため 24 時間モデルより輸送時間が悪化した。しかしながら、平均輸送時間は 14.1[step/BOX](約 7 時間)であり、1 日あたり複数回の幹線輸送が可能であることから、GW モデルでも当日輸送は実現できると考えられる。さらに合計金額やドライバーの負担であるトラック平均移動距離を 24 時間モデルの半分以下に抑えることができたため、GW モデルは金額コストを抑えて 24 時間輸送を実現できることが定量的に評価された。

5 GW の有無による効率の検証

4 章では、全地域に GW を設け、24 時間輸送を全国規模で行った。本章では、GW による 24 時間輸送の地域を限定し、GW モデルと夜間モデルを組み合わせた場合に各評価値がどのように変化するかを検証した。なお、GW の有無を組み合わせた輸送方式を「GW 混合モデル」とする。

5.1 条件設定

GW 混合モデルでは表 1 で設定した GW のうち、取扱い BOX 数の少ない地域から順に GW を廃止し、GW 数を減らしていく。

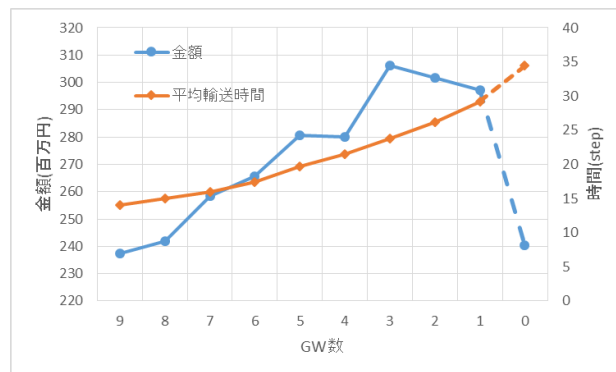


図 4: GW 混合モデルでの平均輸送時間と合計金額

GW のある地域とない地域間の輸送は、GW を有する地域の GW を介して送られるものとする。また、GW のある地域は 24 時間輸送可能、GW のない地域は夜間のみ輸送可能とする。トラックはこれまでと同様、GW 間の輸送のみ 20t トラックを使用する。

5.2 結果

GW の数を減らした際の輸送時間と金額コストの変化を図 4 に示す。なお、図 4 中の GW 数 0 は夜間モデルの結果と同じである。

図 4 からわかるように、GW 数を減らすと平均輸送時間、合計金額ともに高くなった。合計金額については、GW 数が 3 のときが最も高い結果となった。

5.3 考察

平均輸送時間がこのように上昇したのは、GW を廃止することにより夜間まで待機する荷物が増えるためであると考えられる。図 4 より、GW での輸送方式を採用する場合、地域を限定して行うよりも、全国に GW を設置して幹線輸送を行った方が効率がよいことがわかった。

6 結論

6.1 まとめ

シミュレーションにより、夜間のみの幹線輸送方式を 24 時間化した際に発生する問題を定量的に評価し、またその対処法として考えられる GW 方式を用いた場合での各評価値の変化を検証した。その結果、現状の幹線輸送方式をそのまま 24 時間化すると輸送時間が大幅に短くなるが金額コストが大幅に増加することがわ

かった。これは、従来の夜間モデルでは荷物を夜間までためることで集約し、効率的な輸送を行っていたが、24時間モデルでは従来の集約ができないためにより多くのトラックが必要となったためであると考えられる。しかしながら、これらの問題点の対処法として考えられるGW方式では、各地域ごとに荷物を集約して輸送することにより、ドライバーの稼働時間や金額コストを抑えつつ24時間輸送を行うことが可能であることが示された。また、GW方式での輸送の場合、地域を限定してGWを設けるよりも全国で行った方が輸送時間、金額コスト共に良い結果が得られることがわかった。

6.2 今後の展望

本研究では、次の2つの課題が残されている。

- GW配置の最適化

本研究では、各地方の既存の輸送拠点の1つをGWとしてシミュレーションを行った。そのため、GWを新設するとした場合での最適な立地や管轄については現状では考慮できていない。今後は、実際の地理的条件に基づき、GW新設の際の最適な立地を求めた後に各コストの検証を行うことで、より現実に近い値が得られると考えられる。

- 実トラック台数の算出

本研究では、トラックの台数は述べ台数とし、同一トラックの複数回使用については考慮しなかった。しかしながら、一度出発したトラックは、最終的に自身の輸送拠点に戻る必要があり、トラックの復路を考慮した上で金額等のコストを算出することがより現実的であると考えられる。

謝辞

本研究を進めるにあたり、終始多大なるご指導、ご教示を頂きました寺野隆雄教授、吉川厚連携教授に深く感謝致します。また、研究を進める上でのアドバイスだけでなく、論文の書き方や構成など、多くの御助言、御指導をして下さいました山口大学国際総合科学部の山田隆志准教授に深く感謝致します。さらに、物流業務に関してのデータを提供して下さいましたヤマト運輸株式会社様に深く感謝致します。産学連携の共同研究プロジェクトとして現場の知見や現状の問題点など、非常に有意義なご指摘を頂きましたヤマト運輸株式会社村上様に深く感謝致します。また、配車計画に関する基礎研究や最近の動向などで多大なるご協力を頂きました株式会社ライナロジクスの朴成浩社長、西田里華様に深く感謝致します。最後に本研究に当たりご支

援、ご協力を頂きました皆様にこの場を借りて心より御礼申し上げます。

参考文献

- [1] 国土交通省：平成25年度宅配便取扱実績関係資料
- [2] 経済産業省：平成26年度調査結果要旨、電子商取引実態調査
- [3] 林 克彦, 根本敏則：ネット通販時代の宅配便, 成山堂書店 (2015)
- [4] ヤマト運輸株式会社
<http://www.kuronekoyamato.co.jp/top.html>(アクセス 2015/12/04)
- [5] 日本経済新聞：「3度目の革命」ヤマトHD, 2000億円を投じ物流網強化 (2013)
<http://www.nikkei.com/article/DGXNASFK04021v3A700C100000/>(アクセス 2015/11/20)
- [6] 藤杰, 山本学, 寺野隆雄：宅配便ドライバの動的ルートスケジューリング, 計測自動制御学会第5回社会システム部会研究会資料 Vol. 5, pp. 77-84(2014)
- [7] 奥津規矢, 山本学, 寺野隆雄：宅配業務における仕分作業とレイアウト構築の効率化に関する研究, 計測自動制御学会第5回社会システム部会研究会資料 (2014)
- [8] 高橋朋康, 吉川厚, 寺野隆雄：パレットの運用効率を考慮した宅配便配送計画立案に関する研究, 計測自動制御学会第8回社会システム部会研究会資料, pp. 173-178(2015)
- [9] 黒川久幸, 大和裕幸, 小山健夫：ハブ・アンド・スポーク型輸送方式に関する検討, 日本造船学会論文集第184号, pp. 617-625 (1998)
- [10] 徳永幸之, 岡田龍二, 須田熙：宅配輸送におけるセンター配置及び輸送ルート決定モデル, 土木計画学研究・論文集 No.12, pp519-526 (1995)
- [11] V. Pillac, M. Gendreau, C. Guret, and A. Medaglia, "A review of dynamic vehicle routing problems," Eur. J. ..., pp. 028, 2012.
- [12] 株式会社ライナロジクス
<http://lynalogics.com/>(アクセス 2015/12/19)
- [13] 経済産業省：毎月勤労統計調査(平成25年11月)結果確報
- [14] 国土交通省：報道発表資料
「新たな高速料金について」
- [15] 全日本トラック協会：5. 車種区分
<http://www.jta.or.jp/coho/hayawakari/5.kubun.html>(アクセス 2015/12/04)
- [16] NEXCO 中日本：深夜割引
<http://www.c-nexco.co.jp/etc/discount/etc/night/>(アクセス 2015/12/04)
- [17] 国土交通省：自動車燃費一覧, 自動車の燃費基準値
- [18] 経済産業省：資源エネルギー庁, 石油製品価格調査
- [19] 国土交通省：平成22年度交通センサス集計結果整理表

今回発表されない受賞論文の受賞理由

- SIG-SAI 社会におけるAI研究会
「飲食店向け不動産営業を支援する申込み顧客推薦システムの評価」
河村一輝（奈良先端），諏訪博彦（奈良先端），小川祐樹（立命館），荒川豊（奈良先端），
安本慶一（奈良先端）
（SIG-SAI-WSSIT, 第 25 回, 2016 年 3 月 1 日）
選考理由：
本論文は、飲食店向けの不動産営業という具体的な応用に対して、営業マンの電話対応の
ノウハウ不足を補うために、RandomForest を用いたノウハウ推薦の技術を確立し、実デー
タにおいて、53.8%という精度を示すことに成功している。当性能は人間のベテランの精度
に比べても十分に高く、機械学習の実用的応用における有用性を示すことに成功している。
これから、当研究会のテーマである社会での応用を見据えた発表として優秀賞にふさわし
いと考え、これを推薦する。
- SIG-BI：経営課題にAI を! ビジネス・インフォマティクス研究会
「実データに基づく宅配便物流の地域集約方式による効率化の研究」
黒瀬 雄太（東京工業大学），寺野 隆雄（東京工業大学）
（SIG-BI-004-01, 第 4 回, 2016 年 3 月 19 日）
選考理由：
本論文は、宅配便物流に関する課題について、実データに基づいた分析条件のもと、宅配
の輸送方式の輸送時間、コストなど、宅配便物流における主要な課題についてシミュレー
ション手法により取り組んだ社会的価値の高い論文である。数値実験の結果、現在採用さ
れている幹線輸送方式を24 時間化することにより生じる問題点を指摘するとともに、新た
な方式として提案されている地域集約型(GW)方式の有効性を示しており、実用的な研究と
なっている。シミュレーション手法により、ゲートウェイ配置位置に関する数値実験など、
今後の更なる発展に期待できるとともに、実務に密接に関連した課題に取り組んだもので
あり、その有用性の高さから研究会優秀賞に値する。

© 2016 一般社団法人 人工知能学会 合同研究会 2016

〒162 東京都新宿区津久戸町 4-7 OS ビル 402 号室 03-5261-3401 Fax: 03-5261-3402