

戦略的創造研究推進事業 CREST
研究領域「イノベーション創発に資する人工知能基
盤技術の創出と統合化」
研究課題「異種ドメインユーザの行動予測を可能に
するペルソナモデルの転移技術」

研究終了報告書

研究期間 2021年04月～2024年03月

研究代表者：原 隆浩
(大阪大学 大学院情報科学研究科
教授)

§ 1 研究実施の概要

(1) 実施概要

【研究ビジョン】

個人情報の保護に関する法令に則してプライバシーを保護しつつ、ビッグデータ解析・AI 技術により異種サービスを跨ったマーケティング施策を可能とする。特に、利用履歴などのデータが少ない新規サービス事業者や新規ユーザに対しても安心かつ快適なサービスの提供を可能とする。

This project aims to develop big data analysis and AI technologies for digital marketing between different services as well as preserving user privacy in accordance with laws and regulations on protection of personal information. The developed technologies will contribute to provide secure and comfortable services even to new service providers and new users which do not have an enough amount of data including service usage logs.

【達成状況】

3 年間の研究期間中に、多様かつ高度なアプローチの開発とその効果の実データによる検証を行い、想定を超える成果が得られた。学術的な成果として、データベース分野 (VLDB Journal)、Web 分野 (TWEB)、データマイニング分野 (TKDD (3 編))、推薦システム分野 (RecSys (2 編))、人工知能分野 (AAAI) で最高権威の論文誌・国際会議に多数の論文が採択された。現在、ビッグデータ処理、クロスドメイン行動予測に関して世界トップレベルの研究力を有していると確信する。社会的な成果としては、「異業種ペルソナマーケティング AI 推進協議会」(ペルソナ AI 協議会) を設立し、会員企業との協働により、技術の事業応用の可能性や課題を検討している。その一環として、オンライン行動や実空間行動に関する検証実験を実施した。さらに、本研究の成果をもとに 5 件の特許を出願し、実空間におけるペルソナモデリング(エリアモデリング)を実応用に活用するためのベンチャーを起業した。

これらの達成状況を踏まえて、この研究プロジェクトは計画通りに進展し、ビジョンの達成に向けて順調に進んだ。また、当初の計画外の新たな課題にも取り組んでいる。この新課題を含めて、新規サービス事業者や新規ユーザに対する安心かつ快適なサービスの提供に向けたビッグデータ解析・AI 技術の開発とその実用化について、引き続き研究開発を進める予定である。

【概要】

全体概要

本研究では、新規サービス事業者や新規ユーザに対する安心かつ快適なサービスの提供に向けたビッグデータ解析・AI 技術の開発とその実用化について研究開発を進めた。具体的には、オンラインから実空間に至る複数のドメイン(サービス業者やデータ所有者)で構築されるペルソナモデル(ユーザの共通的な行動パターンを表現したモデル)と行動予測モデルを、全く異なる業種のドメインや新規サービス事業者、新規ユーザなどに転移利用するためのクロスドメイン(異種ドメイン)行動予測技術群や、大規模なペルソナモデル(ビッグデータ)を効率的に処理する機能群を開発した。特に、プライバシーや権益の保護に配慮して、運用時に異種ドメイン間でユーザ ID や生データの交換を必要としない、ペルソナモデルや行動予測モデルの転移利用を可能とした。

研究項目 1: 多様なデータソースからのペルソナモデルの構築

オンラインの購買行動ログや、実空間における位置ログ・店舗訪問ログ、SNS や Web などの多様なデータから、ペルソナモデルを構築した。さらに、構築したモデルに対して、状況に応じて効率的かつ効果的にデータ処理を行う機能群を提供した。

1-1: ペルソナモデルの構築

(A) オンライン行動に基づくモデリングとして、E コマースドメインと Web 広告ドメイン、映画・書籍サービスドメインなどのオンラインデータを用いたペルソナモデリングを行った。(B) 空間行動に基づくモデリングでは、公共 WiFi サービスや実店舗購買データを利用し、実空間行動の予測(特定店舗への訪問予測や売上金額予測など)を軸に、これらのデータを用いたペルソ

ナモデリングを行った。さらに、ペルソナモデルの生成につなげるため、大規模なユーザの滞在情報から、場所の使われ方の分散表現(Usage of Area with Stay info.: UAS)を生成する手法(Area2Vec)を提案した。また、(C)環境(外的要因)のモデリングとして、SNS(Twitter)からトレンドやイベント、思想的バイアスなどを検出し、E コマースドメインにおける購買行動など、ユーザの行動予測に活用する技術を開発した。(D)さらに、これまでに構築したペルソナモデルの統合についても研究開発を進めた。(A)は原・小野グループ、(B)は河ログループ、(C)は原グループを中心に実施した。(D)は全グループで実施した。

1-2: ペルソナモデルに対する処理機能群の実現

原グループを中心に、ペルソナモデルなど複雑なデータ構造に対するデータ処理機能群として、空間データからの特徴構造抽出の高速アルゴリズム、内積最大探索に関する高速アルゴリズム、アウトライアが存在するデータ集合から多様な部分集合を抽出する高速アルゴリズム等を開発した。これらの技術は、既存技術と比較して、数十倍以上の高速化を実現した。

研究項目 2: 複数ドメイン間の個人を特定しないペルソナマッピング

原・小野グループで連携し、ID マッチングなどの技術を用いることなく、複数のドメインをつなぐためのユーザやアイテムのマッチングを行う技術の開発を進めた。特に、ドメイン間の特徴の違い(バイアス)を吸収する表現学習の研究開発に取り組んだ。

研究項目 3: ドメイン間の行動予測の転移技術の構築

3-1: 異種ドメインの行動モデルを横断的に扱う統計モデルの構築

原・小野グループを中心に全グループが連携し、異種ドメイン間で行動予測モデルを転移し、ターゲットドメインでの行動予測に活用する技術の開発を行った。また、ユーザとアイテムの関係性を表現可能なデータ構造としてグラフに着目し、異質性を持つグラフのためのグラフ畳み込みネットワークの学習手法や、複数の知識グラフ間で知識補完を行う手法を開発した。

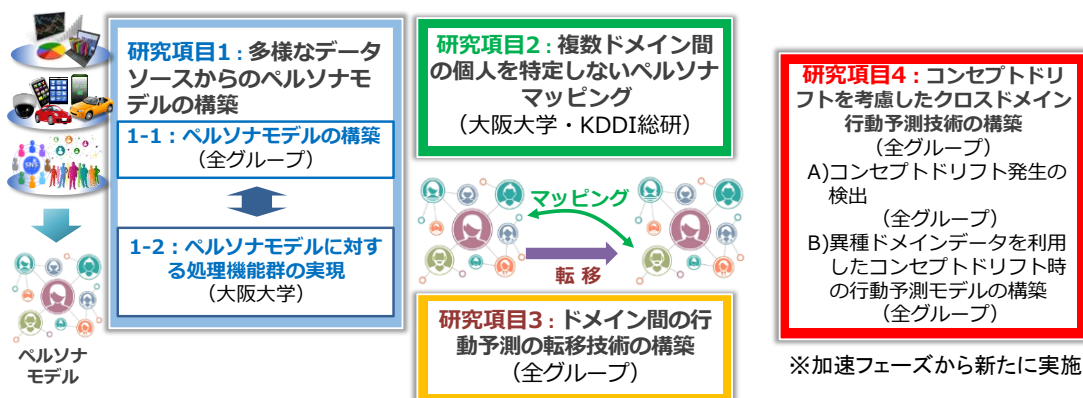
3-2: オンライン行動および実空間行動情報を用いた検証実験

河ログループを中心に、実空間における行動情報を用いてのペルソナモデルの構築に向けて、「ある場所がどのように使われているか」という滞在場所の特徴(UAS)を定義し、様々なエリアでの抽出を通じてその有効性を確認した。また、得られた UAS をクラスタリングし、時間帯や滞在時間で似た特徴を持つ場所を抽出することで、ユーザに「一般にどのように使用されている場所を移動しているか」の情報を付加し、ユーザのモデル化が可能であることを示した。

小野グループを中心に全グループが連携し、2021 年 10 月に設立した異業種ペルソナマーケティング AI 推進協議会での議論を踏まえ、ペルソナを個人情報にならないように符号化し会社間で共有するペルソナ AI サービスを構築することを検討し、実証実験を進めた。

研究項目 4: コロナ禍などの非常時におけるユーザ行動の変動分析

新型コロナウイルスなど非常時の状況下では、オンラインおよび実空間のユーザの行動パターンが大きく変動する。原・小野グループを中心に全グループが連携し、ユーザの行動(インタラクション)の変化等に頑強なグラフベースの行動予測(推薦)モデルなどを開発した。



図：本研究の全体像

(2)顕著な成果

<優れた基礎研究としての成果>

1. ビッグデータ解析・処理技術の革新

概要：研究項目 1-2 として、外れ値検出、空間・キーワードの類似性による結合、逆内積最大探索、位置情報データのコロケーション探索、k 最近傍検索、密度ベースの高速クラスタリング、内積最大探索の結果の多様化、アウトライアを考慮した部分集合抽出などに関して独創的なアプローチおよびそれらの並列処理技術を開発した。これらの技術は、従来技術と比較して 100 倍以上の高速化を実現するなど、ビッグデータ解析・処理技術を革新した。この研究成果は、データベース分野や Web 分野で最高峰の論文誌 VLDB Journal、IEEE Trans. on the Web や、データベース分野、地理情報分野、推薦システム分野、人工知能分野で最難関の国際会議 SIGMOD (2 編)、SIGSPATIAL、RecSys (2 編)、AAAI に採択され、二つの国際会議の最優秀論文賞を受賞するなど、国際的に極めて高い評価を得ている。

2. 環境要因抽出のデータ源としての SNS 活用技術の確立

概要：SNS からトレンドや思考バイアス、イベントなどの環境情報を抽出する技術や、それを推薦システムにおける環境要因として考慮し、推薦精度(行動予測精度)を向上する技術を確立した。これらの技術により、推薦システムの精度が大幅に向上することを確認した。この成果はデータマイニング分野で最高峰の論文誌 ACM Trans. on Knowledge Discovery from Data (2 編)に採択され、国際的に高い評価を得ている。

3. 国際ワークショップの創設

概要：ユビキタス分野でのペルソナモデル学習・転移技術に関する研究分野の確立と発展のため、国際ワークショップ UMUM(Ubiquitous and Multi-domain User Modeling)を創設した。第 1 回は、モバイル・ユビキタス分野で著名な国際会議 EAI MobiQuitous の併設ワークショップとして本チームが主催し、3 件のキーノート/招待講演、5 件の採録論文発表が行われた。第 2 回は、ユビキタス分野で最難関の国際会議 IEEE PerCom の併設ワークショップとして主催し、2 件のオープニングトーク/キーノート、7 件の採録論文発表が行われた。

<科学技術イノベーションに大きく寄与する成果>

1. クロスドメイン行動予測による科学技術イノベーションのためのコンソーシアムの設立

概要：研究成果の社会展開に向けて、連携先募集・技術普及・啓発のための任意団体として、2022 年 10 月に異業種ペルソナマーケティング AI 推進協議会(略称:ペルソナAI協議会)を設立した。これまでに、会員募集のための説明会を複数回開催し、2023 年 10 月現在までに、本チームが所属する組織以外に、9 社が加入した。これらの 9 社は販売業、マーケティング・コンサル業、通信キャリア、ニュースサービス、テレビメディア、ベンチャーなど多種多様な業種で構成されている。この協議会の総会やワーキンググループにおいて、会員企業との情報交換や実証実験、連携の可能性などを議論しており、科学技術イノベーションにつながる基盤を整備できたと考える。

2. クロスドメイン行動予測技術の革新

概要：本研究では、特徴の大きく異なるドメインを想定したクロスドメイン行動予測に関して、運用時(テストフェーズ)に ID マッチングを必要としないプライバシー保護に優れた世界最高レベルの技術と実用化に向けた基盤を築いた。本研究の成果は、新規サービス事業者や新規ユーザに対する安心かつ快適なサービスの提供に向けたビッグデータ解析・AI 技術の開発とその実用化について、今後のイノベーションに大きく寄与すると期待される。

3. 移動データを活用したエリアモデリング、ユーザ(ペルソナ)モデリング技術の確立

概要：従来、位置情報を用いたデータ分析は、機能を把握している地図上にマッピングを行ったり、POI(緯度・経度に対する、機能が説明可能な特定の地物や機能を割り当て)を用いて

分析を行う必要があった。本研究で確立した技術では、シンプルなモデル化とクラスタリングにより、緯度・経度とは独立してエリアの利用形態を類型化することを可能とする。さらに、このエリアモデリングに基づき、その間の移動経路 (Trajectory) をモデル化することが可能となる。これらの技術は、実店舗の来店予測や購買予測などのマーケティングだけではなく、人流予測に基づく都市計画など、公共サービスのイノベーションに大きく寄与すると期待される。

< 代表的な論文 >

1. Yihong Zhang, Takahiro Hara, Explainable Integration of Social Media Background in a Dynamic Neural Recommender, ACM Transactions on Knowledge Discovery from Data, Vol.17, No.3, pp.1-14, 2023.

概要：この論文では、ニューラルネットワークに基づく推薦システムに SNS を活用するための説明可能な手法を、世界で初めて考案している。この手法では、相関ルールマイニングを活用し、SNS と E コマースサービスにおけるユーザの行動パターンを人が理解可能な形式で抽出する。抽出した行動パターンを用いた学習モデルを構築し、SNS と E コマースドメインにおけるユーザの行動予測や推薦を実行する。評価実験により、考案手法は、既存技術と比較して、大幅に高い精度を示すことを確認した。

2. Daichi Amagata, Takahiro Hara, Reverse Maximum Inner Product Search: Formulation, Algorithms, and Analysis, ACM Transactions on the WEB, Vol.17, No.4 Article 26, 2023.

概要：逆内積最大探索という新しい問題を数学的に定義し、それを高速に実行するアルゴリズムを考案し、性能解析の結果を示した論文である。この問題は、例えば商品の推薦において、対象の商品に興味をもつユーザ群を高速に発見することができるため、推薦システムにおいて重要な基盤操作となる。評価実験により、考案したアルゴリズムは、既存技術と比較して、数百倍高速に動作することを示した。

3. Zhi Li, Daichi Amagata, Yihong Zhang, Takahiro Hara, Shuichiro Haruta, Kei Yonekawa, Mori Kurokawa, Semantic Relation Transfer for Non-overlapped Cross-domain Recommendations, Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2023.

概要：共通ユーザや共通アイテムが存在しない異種ドメイン間において、意味的關係を転移しクロスドメイン推薦を実現することを目指した論文である。この論文の考案手法では、両ドメインのアイテムに意味的クラスタリングを施し、両ドメインを紐づける。さらに、各クラスタにおけるアイテム間の意味的關係の類似性を調べることでドメイン間の意味的差異を検出し、両ドメイン間の特徴の違いを考慮した学習モデルを構築する。評価実験により、考案手法が既存のクロスドメイン推薦手法よりも大幅に高い推薦精度を示すことを確認した。

§ 2 研究実施体制

(1) 研究チームの体制について

① 原グループ

研究代表者：原 隆浩（大阪大学大学院情報科学研究科・教授）

研究項目

- ・ ペルソナモデルの構築 (オンライン行動に基づくモデリング)、ペルソナモデルの構築 (環境モデリング)、ペルソナモデルの構築 (統合)、ペルソナモデルに対する処理機能群の実現
- ・ データの特性に適応したペルソナマッピング、ユーザグループのマッピング
- ・ 異種ドメインの行動モデルを横断的に扱う統計モデルの構築、ペルソナ間社会的ネットワークの推定技術、オンライン行動および実空間行動情報を用いた検証実験
- ・ コンセプトドリフトを考慮したクロスドメイン行動予測技術の構築

② 河ログループ

主たる共同研究者：河ロ 信夫（名古屋大学・大学院工学研究科・教授）

研究項目

- ・ ペルソナモデルの構築(オンライン行動に基づくモデリング)、ペルソナモデルの構築(環境モデリング)、ペルソナモデルの構築(統合)
- ・ 異種ドメインの行動モデルを横断的に扱う統計モデルの構築、ペルソナ間社会的ネットワークの推定技術、オンライン行動および実空間行動情報を用いた検証実験
- ・ コンセプトドリフトを考慮したクロスドメイン行動予測技術の構築

③ 小野グループ

主たる共同研究者: 小野 智弘 (KDDI 総合研究所・Human-Centered AI 研究所・所長)

研究項目

- ・ ペルソナモデルの構築(オンライン行動に基づくモデリング)、ペルソナモデルの構築(環境モデリング)、ペルソナモデルの構築(統合)
- ・ データの特性に適応したペルソナマッピング、ユーザグループのマッピング
- ・ 異種ドメインの行動モデルを横断的に扱う統計モデルの構築、ペルソナ間社会的ネットワークの推定技術、オンライン行動および実空間行動情報を用いた検証実験
- ・ コンセプトドリフトを考慮したクロスドメイン行動予測技術の構築

(2)国内外の研究者や産業界等との連携によるネットワーク形成の状況について

本研究チームは、研究期間中に国際ワークショップ Workshop on Ubiquitous and Multi-domain User Modeling (UMUM)を創設し、2度開催した。その中で、行動認識分野で第一人者のジョージア工科大学 Thomas Ploetz 准教授を招待講演者として招くなど、関連分野の研究者との国際的なネットワーク形成に努めた。研究代表者の原は、長年に亘り、データベース分野や知識処理分野で国際的に活躍しており、国際共同研究の経験も豊富である。さらに原は、2023年度に開始した ACT-X「次世代 AI を築く数理・情報科学の革新」の研究総括を務め、AIP ネットワークラボの構成領域のメンバーとしても参画しており、国内外の様々な研究者と交流できる環境にある。主たる研究分担者の河口は、代表を務める人間行動センシングコンソーシアム (HASC) を通じて世界最大規模の行動データベースを収集・公開すると共に、国際ワークショップ HASCA を 2013 年から連続で開催している。小野は、積極的な国際連携を進め、2020 年以降では継続学習の第一人者のイリノイ大学シカゴ校 Bing Liu 教授、転移学習の第一人者の南カリフォルニア大学 Yan Liu 教授、GNN の第一人者スタンフォード大学 Jure Leskovec 教授など 8 名の著名教授陣とそれぞれ共同研究を立ち上げるとともに、人工知能学会全国大会での Bing Liu 教授を招聘した企画セッションを実施するなど今後のさらなる連携の準備を進めている。このように、本研究チームは、クロスドメイン行動認識に関する国内外での交流や共同研究を通じて、新たな研究コミュニティの形成を進めている。

産業界との連携については、社会実装の検討やその前段となる実証実験の企画、課題抽出を進めることを目的とし、2021 年 10 月に設立した異業種ペルソナマーケティング AI 推進協議会を設立した。これまでに、総会 4 回、シンポジウム 1 回を開催した。本協議会には、2023 年 10 月現在までに、本チームの所属組織以外に 9 社が入会している。これらの協議会活動の一環として、会員企業と共に本研究における実証実験①～③を完了した。2022 年度は、①オンライン行動の実証 1 件を完了し、特定のニュースカテゴリに対するニュース推薦において既存手法に対し 0.15%のクリック率向上を確認した。さらに 2023 年度は、②オンライン行動に関して、異なる会員企業での閲覧記事、購買商品のデータに対し大阪大学(KDDI 総合研究所共同)開発のグラフ知識転移手法を適用し、特定の商品に対し購入意向が 1.3 倍高い層を特定するという効果を確認した。また③実空間行動に関して、レシート情報で得られる購買傾向を反映した Ormis・名古屋大学開発のエリア・ユーザモデリングの結果、在住地域により購買傾向が異なるインサイトを獲得し、レシート情報と位置情報という異種ドメインの掛け合わせによるモデリングの有用性を確認した。このように、研究成果を社会還元するためのネットワーク構築は十分に進んでいる。