

オンラインカンファレンスにおける情報推薦導入の検討

大社 綾乃[†] 大滝 啓介[†] 石井 良尚[†] 中島 健介[†] 川崎 滉平[†]
小出 智士[†]

[†] 株式会社豊田中央研究所 〒480-1118 愛知県長久手市横道 41-1
E-mail: †{okoso,otaki,y-ishii,e1403,e1734,koide}@mosk.tytlabs.co.jp

あらまし 新型コロナウイルス流行の影響により、多くの学術会議や展示会がオンラインで開催されている。本稿ではこれらを総称してオンラインカンファレンスと記す。カンファレンスにおいて参加者が高い満足度を得るためには、自身の興味のある展示物を素早く見つけ情報収集すること（適合性）や、意外性のある展示物を発見し新たな分野を開拓すること（セレンディピティ）が重要な要素だと考えられる。そこで本研究では、オンラインカンファレンスにおける適合性およびセレンディピティの向上を目指し、情報推薦の導入を検討する。具体的には、オンラインカンファレンスのログデータを用いてユーザの行動を分析し、オフライン評価によって情報推薦が有効であるかを検証する。

キーワード オンラインカンファレンス、情報推薦、協調フィルタリング、行列分解

1 はじめに

新型コロナウイルス流行の影響やオンライン環境の拡充により、多くの学術会議や研究会がオンラインで開催されている。会議の形式には様々なものがあり、例えば zoom¹ などの web 会議システムを用いてリアルタイムで研究発表を行うものや、web サイト上に研究発表動画を掲載するものがある。本稿では、これらを総称して「オンラインカンファレンス」と記す。オンラインカンファレンスは、参加人数に上限がなく世界中のどこからでも参加でき、開催コストも小さい。将来のカンファレンスはオンサイトとオンラインの欠点を補完し合うハイブリッド開催が主流になっていくだろうと言われている [3], [8]。

カンファレンスの参加者が高い満足度を得るためには、以下の3つの体験が重要だと考えられる。

適合性：自身の興味のある展示物を素早く見つけ情報収集する
セレンディピティ² [5], [9]：意外性のある展示物を発見し新たな分野を開拓する

人脈形成：人脈を形成しビジネスパートナーを見つける

参加者は自身に興味のある展示物であるかを主にタイトルや概要文から判断する必要があるため、大量にある展示物の中から自身の興味に適合した展示物をどのようにして素早く見つけるかが重要である。また、オンラインではオンサイトと比較してセレンディピティの高い体験が難しいため、オンラインカンファレンスにおけるセレンディピティ向上は重要な課題である。例えばオンサイトカンファレンスの参加者は、自身の興味のある研究発表と同セッションの他の発表や、興味のあるポスター発表の周辺のポスターなどが自然と目に入り、聴講を予定

していなかった研究発表と出会えることが多くある。オンラインカンファレンスでは、聴講予定の研究発表以外のものが目に入る機会が少なく、このような新たな出会いは少ないと考えられる。オンラインカンファレンスにおける人脈形成も重要なトピックであるが、本研究で扱うデータでは分析できないため、本稿では**適合性とセレンディピティ**に着目する。

オンラインカンファレンスにおける適合性およびセレンディピティの向上を目指し、我々は情報推薦の導入を検討する。情報推薦とは、ユーザの要求を推定し、その要求を満たす情報を提示するモデルである [2], [10]。Amazon³ などの EC サイトや Netflix⁴ に代表される動画配信サービスにおいて、ユーザが興味を持ちそうな商品や動画を提示する機能として活用されている。適合性やセレンディピティは情報推薦の分野で鍵となる指標であり、情報推薦技術を導入することで、適合性やセレンディピティの向上が期待できる。現在オンラインカンファレンスが多くの開催されている一方で、情報推薦技術を取り入れているものは我々の知る限りほとんどない。

本稿では、オンラインカンファレンスにおいて情報推薦が有効かどうかを検討する。具体的には以下の Research Question に取り組む。

- **RQ1 (適合性)**：ユーザはどのように自身の興味のある展示物を発見しているのか。
- **RQ2 (セレンディピティ)**：ユーザは自身の興味と関連のない（潜在的に興味のある）展示物も閲覧しているのか。
- **RQ3 (有効性)**：オンラインカンファレンスにおいて情報推薦技術は有効であるのか。

これらの RQ に答えるため、2021 年 10 月 12 日～25 日にオンラインで開催された、グループ会社向け研究発表会のデータを用いてユーザの行動を分析し、オフライン評価により情報推薦

1：Zoom Video Communications, Inc. (<https://explore.zoom.us/ja/products/meetings/>)

2：目新しさや意外性を意味する。

3：Amazon.com, Inc. (<https://www.amazon.co.jp/>)

4：Netflix, Inc. (<https://www.netflix.com/jp/>)

技術の有効性を検証する。本研究の貢献は以下の通りである。

(1) オンラインカンファレンスでのユーザの行動ログを取得し、その統計情報およびユーザの行動を分析した結果をまとめた (3・4 章)。

(2) 古典的な協調フィルタリング手法である行列分解を用いて、オンラインカンファレンスにおいて情報推薦技術が有効であることをオフライン評価により示した (5 章)。

本稿の構成は次の通りである。2 章では本稿の前提となる定義や知識について述べる。3 章では分析に用いたオンラインカンファレンスのデータについて概説する。4 章ではオンラインカンファレンスにおけるユーザの行動を分析し、**RQ1 (適合性)** および **RQ2 (セレンディピティ)** について議論する。5 章ではオフライン評価を行い、**RQ3 (有効性)** を議論する。6 章にて本稿をまとめ、今後の課題について論ずる。

2 準備

本章では、本稿で扱うオンラインカンファレンスの構成および記録されるユーザの行動を定義する。また、セレンディピティ向上に有用な協調フィルタリング、および 5 章の実験で用いた非負値行列因子分解について概説する。

2.1 定義

オンラインカンファレンスの web サイトはページ集合 P から成り、各ページ ID を $p \in P$ とする。web サイトを閲覧できるユーザ集合を U 、ユーザ ID を $u \in U$ とする。オンラインカンファレンスは $|N|$ 日間開催され、開催日を $N := \{1, \dots, n, \dots, |N|\}$ とした時、 n 日目に web サイトを訪問するユーザ集合を U_n とする。ここで $|N|$ は集合 N の要素数とする。ユーザは web サイトを訪問すると初めにトップページを閲覧する。その後はページ集合 P 内のページ間を遷移し、各展示物 $e \in E$ を閲覧する。 e は展示物の ID、 E は展示物の集合を表し、展示物のカテゴリ集合を C とした時、各展示物はいずれかのカテゴリ $c \in C$ に属する。なお、カテゴリ c の展示物集合を E_c と表す。展示物が掲載されたページ集合を $P_E \subset P$ と表し、すべての展示物 e は P_E のいずれかのページに一一対一対応で掲載される。全ての展示物には動画およびイイねボタンが掲載されている。

ユーザの行動ログ集合を L 、ユーザ u の行動ログを $L_u \subseteq L$ とする。各ログは $l := (u, n, t, a, p) \in L_u$ で表し、 u はユーザ ID、 n は開催日、 t は時刻、 $a \in A$ はアクション、 p はページ ID を意味する。 $A := \{a_c, a_m, a_l, a_q\}$ はアクション集合を表し、 a_c はページの遷移または更新、 a_m は動画視聴開始ボタンのクリック、 a_l はイイねボタンのクリック、 a_q は展示物へのアンケート回答を意味する。ログ L のうちでアクションが a_c であるものを $L_c = \{(u, n, t, a, p) \mid a = a_c\}$ とする。同様に L_m, L_l および L_q を定義する。

2.2 古典的な推薦手法

推薦を受けるユーザ (以降は対象ユーザと記す) の嗜好を推定する古典的な手法は、内容ベースフィルタリングと協調フィルタリングに大別される [1]。内容ベースフィルタリングではアイ

テムの特徴ベクトルと対象ユーザの嗜好を比較し、対象ユーザが好むと予想されるアイテムを推薦する。アイテムと対象ユーザのデータのみを用いるため、推薦システムを利用するユーザ数が少なくても利用可能であるが、類似したアイテムが推薦されやすいといった欠点もある。協調フィルタリングではアイテムへの評価が類似しているユーザは嗜好が類似しているという直感に基づき、他のユーザの評価情報を用いてアイテムを推薦する。他のユーザ情報を用いることで、対象ユーザ自身が知らない特徴を持つアイテムを知ることができるが、他のユーザによる評価が十分でないアイテムの推薦が難しいというコールドスタート問題がある。本稿ではセレンディピティに着目するため、他のユーザ情報を用いて対象ユーザ自身が知らない特徴を持つアイテムを推薦可能な協調フィルタリングに焦点を当てる。

2.3 非負値行列因子分解

協調フィルタリングの 1 種として広く推薦システムで用いられている手法として、行列分解がある [4]。行列分解は、要素 (i, j) がユーザ i のアイテム j への評価値からなる評価値行列があるとき、評価値行列内の評価済みの値から未評価値を予測するための手法として用いられる。行列分解の中でも、非負値の要素のみからなる低ランク行列に分解する手法を非負値行列因子分解 (Nonnegative Matrix Factorization; NMF) という [6]。NMF は $m \times n$ の評価値行列 \mathbf{R} を以下のように $m \times k$ のユーザ行列 \mathbf{U} と $n \times k$ のアイテム行列 \mathbf{V} に近似分解する。

$$\mathbf{R} \approx \mathbf{U}\mathbf{V}^\top \quad \text{s.t. } \mathbf{U}, \mathbf{V} \geq 0 \quad (1)$$

ここで \top は転置を表し、 $k \ll \min(m, n)$ である。ユーザ i のアイテム j への評価値 \hat{r}_{ij} は、 \mathbf{U} の第 i 行ベクトル \mathbf{u}_i と、 \mathbf{V} の第 j 列ベクトル \mathbf{v}_j の内積 $\hat{r}_{ij} = \mathbf{u}_i^\top \mathbf{v}_j$ で求めることができる。非負値行列因子分解では交互最小二乗法や勾配降下法を用いて式 (1) で近似した行列の損失を減らす操作を行う。 Ω は評価済みのデータ集合とした時、具体的には以下の式を最小化する。

$$\sum_{(i,j) \in \Omega} (r_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2 \quad (2)$$

各ユーザが評価したアイテムが全アイテムの中で極端に少ない場合に、行列分解は有効な手段であると知られている。

3 データ

本章では、2021 年 10 月 12 日~25 日に開催されたオンラインカンファレンスにおける、web サイトの構成および取得されたユーザのログデータについて概要を述べる。本カンファレンスは、グループ会社向けに研究発表を行うことを目的としている。なお、カンファレンスにおける発表タイトルやカテゴリ名については機密の関係上、本稿では匿名化して扱っている。取得されたデータの詳細は表 1 にまとめる。

3.1 web サイトの構成

オンラインカンファレンスの web サイトのページ構成を図 1 に示す。TOP ページは全てのユーザが一番最初に必ず訪れる

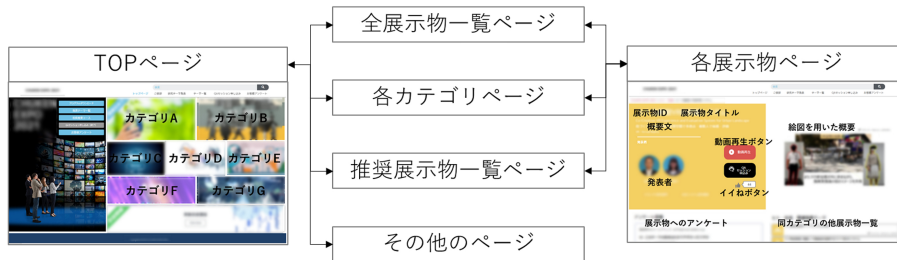


図 1: web サイトのページ構成

表 1: データの詳細

ページ数 $ P $	ユーザ数 $ U $	展示物数 $ E $
53	4133	40
カテゴリ数 $ C $	ログ総数 $ L $	ページ遷移・更新数 $ L_c $
7	133982	96975
動画視聴数 $ L_m $	イイね数 $ L_l $	アンケート回答数 $ L_q $
30562	1396	5049

表 2: ユーザの行動に関する統計情報の詳細

	平均値	中央値
訪問カテゴリ数	3.3	3.0
訪問展示物数	6.8	5.0
動画視聴数	5.9	4.0
動画視聴時間の割合	0.86	1.0
イイねクリック数	0.37	0.0
アンケート回答数	1.2	0.0

ページである。TOP ページ上には各展示物のカテゴリが表示され、それぞれのボタンをクリックすることで、各カテゴリのページへと遷移できる。全展示物一覧ページでは、各カテゴリごとに全ての展示物が表示され、直接展示物ページへと遷移できる。推奨展示物一覧ページでは、視聴を推奨する展示物を表示し、直接展示物ページへと遷移できる。カテゴリページは各カテゴリに関する概要文と絵図、各展示物へと遷移するボタンが表示される。展示物ページでは、各展示物の概要文および絵図、同じカテゴリの別展示物へ遷移するボタン、アンケートが表示される。全ての展示物ページには、それらを紹介する動画再生ボタン、ユーザが良かったと思ったかどうかを記録する「イイね」ボタンがあり、イイね数はリアルタイムで公表されている。一部のカテゴリページ ($c \in \{A, B, F\}$) のみ、カテゴリを紹介する動画とイイねボタンがある。その他のページではオンラインカンファレンス全体に対するアンケートページなどがあるが、本稿では用いないため詳細な説明は割愛する。本オンラインカンファレンスでは、カテゴリ数 $|C| = 7$ 、展示物数 $|E| = 40$ 、ページ総数 $|P| = 53$ であった。なお、各カテゴリに含まれる展示数は $|E_A| = 10$ 、 $|E_B| = 8$ 、 $|E_C| = 5$ 、 $|E_D| = 5$ 、 $|E_E| = 5$ 、 $|E_F| = 3$ 、 $|E_G| = 4$ であった。

3.2 ユーザデータ

ユーザは事前登録することにより web サイトを閲覧できる。事前登録時には、名前、所属会社、所属部署、役職、メールアドレスを入力し、一意のユーザ ID が付与されて記録される。なお、事前登録時にユーザデータおよびサイト訪問ログを個人が特定できない形で研究利用することに同意を得ている。具体的には、ユーザ ID、所属会社、所属部署のみを記録対象としたが、本稿ではユーザ ID のみを扱っているため詳細な説明は割愛する。事前のユーザ登録数は 5591 名であったが、実際にサイトを訪れたのは 74% に相当する 4133 名であった。以降は $|U| = 4133$ として検証した。

3.3 ログデータ

行動ログの総数は $|L| = 133982$ で、ページの遷移または更新数は $|L_c| = 96975$ 、動画視聴数は $|L_m| = 30562$ 、イイねボタンのクリック数は $|L_l| = 1396$ 、展示物へのアンケート回答数は $|L_q| = 5049$ であった。何らかの理由によりユーザ ID が欠損しているデータは除き、同一ユーザが同じ動画に対して動画視聴開始ボタンを連続して複数回クリックしていた場合、一番最初のログのみを残している。また、同一ユーザが同じアンケートに複数回答している場合や、同じ展示物のイイねボタンを複数回クリックしている場合は、一番最後のログのみを残す処理を施している。

4 データ分析

本章では、RQ1 (適合性) および RQ2 (セレンディピティ) について考察するため、取得されたユーザのログデータの分析について議論する。まず始めに基本的な統計情報について述べた後、ユーザが展示物への興味の度合いを示す動画視聴に関する分析について述べる。その後、RQ1 (適合性) および RQ2 (セレンディピティ) について議論する。

4.1 基本的な統計情報

本節ではユーザの行動および展示物に関する統計情報について述べる。なお、以下の統計情報を根拠づける図および詳細は付録に記載している。

a) ユーザの行動に関する統計情報

ユーザの行動に関する統計情報の詳細を表 2 にまとめる。75% のユーザが複数カテゴリのページを遷移し、平均カテゴリ訪問数は 3.3 であった。訪問した平均展示物数は 6.8 で、中央値は 5.0 であった。ユーザが展示物へ実際に興味があったかを示すインタラクションは、動画視聴・イイねボタンのクリック・アンケートへの回答の 3 種類ある。平均動画視聴数は 5.9 本で

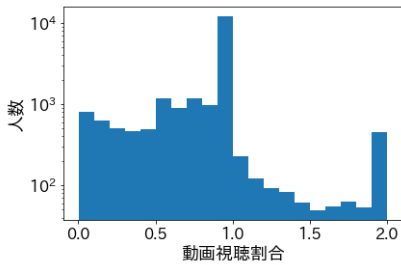


図 2: 動画視聴時間の割合

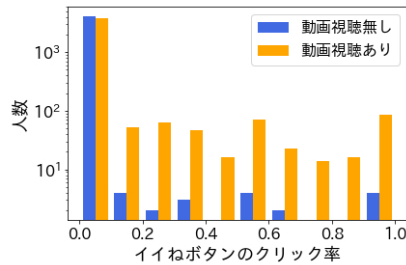


図 3: 動画視聴有無とイイね率

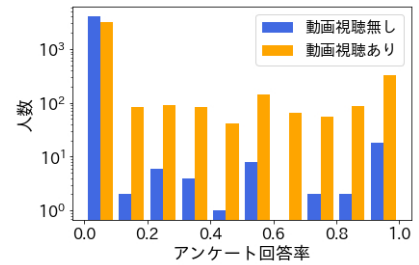


図 4: 動画視聴有無とアンケート回答率

あったが、ユーザの 88%が一度もイイねボタンをクリックせず、75%は一度もアンケートに回答していなかった。そのため以降の分析では、ユーザの興味の有無を示す指標として動画視聴に関する情報を用いる。

b) 展示物に関する観察

カテゴリ A への訪問者数が最も多く、ユーザの興味が高さが窺えた。訪問者数が多かった要因として、展示物数が多かったことや、TOP ページにおけるカテゴリ A への遷移ボタンが目につきやすい上部に配置されていたことが挙げられる。各展示物について、訪問者数に対する動画視聴率には大きな差はなく、60%~80%の割合で動画が視聴されていた。

4.2 動画視聴に関する分析

本節では、動画視聴に関する情報をユーザの展示物に対する興味の度合いとして用いるための前準備として、以下の 2 点について分析した結果を述べる。

- (1) 各動画をどの程度視聴しているのか。
- (2) 動画を視聴した場合としない場合で、ユーザの振る舞い(イイねクリック率やアンケート回答率)は変化するか。

まず始めに、ユーザが各動画をどの程度視聴しているかを調査した。ユーザの行動ログデータから、各動画の視聴時間を計算した。具体的には、ある開催日 n の各ユーザ $u \in U_n$ のログデータ L_u を時刻順に並び替え、動画視聴開始ボタンのクリックログから次のログまでの時間を動画視聴時間とした。1 回の動画視聴時間がその動画の実時間より長い場合には、動画視聴時間を動画の実時間とした。各展示物に対して、各ユーザの動画視聴時間の合計を動画の実時間で割った値を動画視聴時間の割合とする。動画視聴時間の割合を示すヒストグラムを図 2 に示す。縦軸は対数で人数を表し、横軸は動画視聴時間の割合を表す。なお、動画視聴時間の割合が 2.0 より大きい場合には 2.0 とした。図 2 より、多くのユーザが動画を実時間分視聴し、一部のユーザは 2 回以上視聴していることが分かった。動画視聴時間の割合の平均値は 0.86、中央値は 1.0 であった。

次に動画視聴後のユーザの振る舞いについて調査した。訪問した展示物に対して、動画視聴した場合としない場合のイイねボタンのクリック率を比較したヒストグラムを図 3 に、同様の場合のアンケート回答率を比較したヒストグラムを図 4 に示す。横軸は割合、縦軸は対数で人数を表す。イイねボタンのクリック、アンケートの回答の両方において、動画を視聴した場合の方がその振る舞いが増加していた。このことから、動画

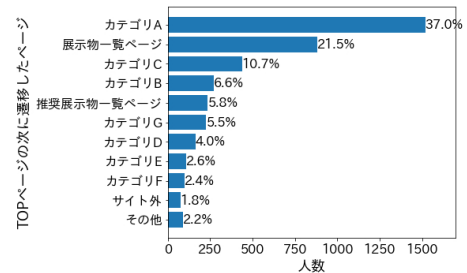


図 5: TOP ページの次に遷移したページ

視聴に関する情報はユーザの興味の度合いを十分に表していると言えるため、以降の検証では動画視聴と展示物に対する興味の度合いに対して以下のように仮定して扱う。

動画視聴が無い場合 興味に適合していない展示物、
動画視聴時間の割合が十分でない場合 興味のある展示物だろうと予想して動画を視聴し始めたが、実際には自身の興味とは適合していないと判断した展示物、
動画視聴時間の割合が十分である場合 興味のある展示物だろうと予想し、さらに実際に自身の興味と適合していた展示物。
 なお、ここで動画視聴時間の割合が十分であるかどうかは、動画視聴時間の割合が閾値 w 以上であるかで判定する。

4.3 RQ1 (適合性)

RQ1 (適合性) へ回答するため、ユーザがどのように自身の興味のある展示物へとページ遷移したかを調査した。初訪問時の TOP ページの次に遷移したページを比較した結果を図 5 に示す。各カテゴリページへと遷移したのが 69%、展示物一覧ページへが 22%、推奨展示物一覧ページへが 5.8%であった。TOP ページの上部に遷移ボタンが配置されていたカテゴリ A へのページ遷移が最も多く、37%であった。また推奨展示物一覧ページへ遷移した 5.8%のユーザは、閲覧したい展示物がなく、お薦めから選んだ可能性があると考えられる。1.8%の 74 名のユーザは TOP ページのみを訪問し、他のページへは訪問しなかった。以上の結果より、**自身の興味のある展示物を階層的に探索しているユーザが多い**と結論づけられる。

4.4 RQ2 (セレンディピティ)

RQ2 (セレンディピティ) へ回答するため、「ユーザが自身の興味と関連のない(潜在的に興味のある)展示物を閲覧しているか」を調査する。本オンラインカンファレンスは各展示物

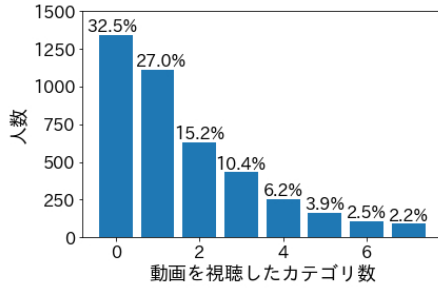


図 6: 十分に動画を視聴したカテゴリ数 ($w = 0.5$)

の専門性が高いことを考慮し、興味のある展示物は同一カテゴリ内である可能性が高いと仮定する。すなわち、「複数カテゴリの展示物の動画を十分視聴している場合には、ユーザが潜在的に興味のある展示物を探索している (=セレンディピティがある)」と仮定して検証する。以降では、どの程度のカテゴリ数にわたって展示物を閲覧しているのか、おおよびどのような分布で複数カテゴリの展示物を閲覧しているのかを分析し、**RQ2 (セレンディピティ)** について議論する。

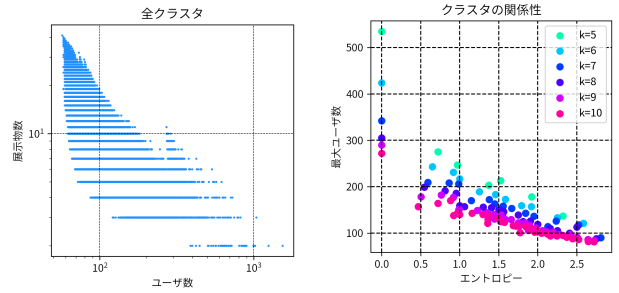
4.4.1 動画を視聴したカテゴリ数

どの程度のカテゴリ数にわたって展示物を閲覧しているのかを調査する。自身の興味に適合するかどうか探索しながら展示物を訪問したのではなく、ユーザが自身の興味と適合した展示物を閲覧したカテゴリを抽出するため、動画視聴時間の割合が閾値 w 以上であったものを十分に動画を視聴したと判定して、そのカテゴリ数を算出した。動画視聴開始ボタンのクリックが 1 回以上あったユーザを対象に、十分に動画を視聴したカテゴリ数を図 6 に示す。なお閾値を $w = 0.5$ とし、横軸はカテゴリ数、縦軸は人数を表す。図 6 より、33%のユーザは十分に展示物を閲覧していないため、興味のある展示物だと思って動画を視聴し始めた展示物が、実際には自分の興味とは適合していないと判断したと予想される。27%のユーザは同一カテゴリしか閲覧しておらず、興味と適合していたカテゴリは 1 つのみで、セレンディピティが無かったと言える。また、動画を視聴したカテゴリ数が 6 以上のユーザは、セレンディピティがあったのではなく、ほぼ全てのカテゴリの展示物を網羅的に探索したのだと考えられる。動画を視聴したカテゴリ数が 2~5 の 35.7%のユーザは、セレンディピティのある体験ができたと予想されるが、網羅的に探索したユーザとの境目については詳細な解析の余地がある。以上の結果より、**27%のユーザは潜在的に興味のある展示物の閲覧はなかった**と結論づけられる。

4.4.2 クラスタ分析

ユーザが訪問した展示物についてより詳細に解析するために、展示物のクラスタ分析を行う。以下のようにユーザ × 展示物のペアの訪問有無を表す行列 \mathbf{Y} を作成する。

$$\mathbf{Y} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1|E|} \\ r_{21} & r_{ij} & \dots & r_{2|E|} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{m|E|} \end{pmatrix} \quad (3)$$



(a) 全クラスタの分布 (両対数) (b) エントロピーと $|U|$ の分布

図 7: 訪問履歴の解析結果

ここで $m = |U|$ であり、ユーザ i が展示物 j を訪問した場合は $r_{ij} = 1$ 、訪問しなかった場合は $r_{ij} = 0$ となる。行列 \mathbf{Y} から構築した二部グラフを作成し、極大二部クリークを列挙する。極大二部クリークを列挙するために、トランザクションデータからの頻出閉集合発見手法 [7] を用いた。なお本データは [7] で研究されているデータよりも密である。そのため多数のクラスタが出力されると予想され、具体的に \mathbf{Y} から極大二部クリーク列挙を行った結果、12,582,231 個のクラスタを得た。各クラスタ (U', E'), $U' \subseteq U, E' \subseteq E$ について、ユーザ数 $|U'|$ と展示物数 $|E'|$ を散布図としてプロットしたものを図 7(a) に示す。大まかには顧客数は展示数に反比例する。一方で、図 7(a) で飛び出している箇所のように、一部のクラスタにおいて多くの展示物を見る傾向が見られた。

より詳細に考察するため、訪問展示物数が $5 \leq |E'| \leq 10$ の範囲を抽出して観察した。展示物 $e \in E'$ に紐付いたカテゴリ ID を利用し、集合 E' が均一なカテゴリを含むのか、多様なカテゴリを含むのかをエントロピーで計測した。また図 7(a) の訪問展示物数とエントロピー値ごとに、最大のユーザ数を持つクラスタを抽出した散布図を図 7(b) に示す。結果より、いずれの場合でも 1 つのカテゴリしか視聴しないユーザが一定数存在すること (横軸が 0.0 の部分に相当)、訪問展示物数の増加に伴いクラスタのユーザ数も減少すること、**多様な展示物を訪問するユーザは均一な展示物を訪問するユーザと比較して少数である**こと (図 7(b) が大まかに右肩下がりになる) を確認した。

4.5 まとめ

上記の分析結果をもとに、**RQ1 (適合性)** および **RQ2 (セレンディピティ)** についてまとめを述べる。ユーザの 69% が自身の興味のある展示物を階層的に探索していた (**RQ1 (適合性)**)。web サイトの構成上、階層的に展示物を閲覧した場合には他のカテゴリの展示物が目に入らないため、セレンディピティが起きづらいと考えられる。実際に動画を視聴したユーザのうちの 27% は同一カテゴリしか十分に動画を視聴しておらず、セレンディピティのある体験はできなかったと言える (**RQ2 (セレンディピティ)**)。オンラインカンファレンスを対象とした推薦技術によってセレンディピティを生み出し、ユーザの満足度を高めることが可能であると期待できる。

5 オフライン評価

本章では **RQ3 (有効性)** を検証するために行ったオフライン評価について述べる。「ユーザがいくつかの展示物を訪問した際に、その訪問履歴をもとにお勧めの展示物を提示する」というシナリオを考える。この提示により展示物の訪問率や動画視聴率が向上すれば、情報推薦が有効であると言える。しかし本研究ではオンライン評価は行っていないため、オフライン評価で検証可能な「ユーザが時系列順で x 箇所の展示物の動画を視聴した際、残りの展示物の動画を十分視聴するか（動画の視聴時間の割合が $w > 0.5$ であるか）否かの二値分類を行う」というタスクに置き換える。このタスクの推定精度が高い場合に、情報推薦の有効性に展望があるとみなせる。ただし、実際に情報推薦がオンラインカンファレンスにおいて有効であるかの検証にはオンライン評価が必要であるため、本評価は予備検証の立ち位置であることを付記しておく。

実際のオンラインカンファレンスにおける推薦システムを想定すると、 n 日目のユーザ $u \in U_n$ へ推薦する展示物を予測する際、 $n-1$ 日目までの全ユーザのデータを用いることができる。ただしコールドスタート問題を考慮して、 $n \geq 2$ とする。上述したタスクにおいても、 n 日目のユーザ $u \in U_n$ について予測する際には、 $n-1$ 日目までの全ユーザのデータを用いる。本評価ではユーザ数の少ない土日を除き、10 日間を対象とした。

5.1 評価指標

本節ではオフライン評価のための評価指標について述べる。次節で説明する予測モデルは、 x 個の展示物の動画を視聴済みのユーザに対して、残りの $|E| - x$ 個の展示物をスコアリングしておすすめ順に並び替えるモデルである。このスコアリングの良さを評価するため、本研究では ROC 曲線 (Receiver Operating Characteristic curve) の AUC (Area Under the Curve) を用いる。ROC 曲線とは、閾値 θ を変化させた時の分類結果の真陽性率 (True Positive Rate : TPR) を縦軸に、偽陽性率 (False Positive Rate : FPR) を横軸にプロットした曲線である。なお分類結果は表 3 に示す通り 4 種類あり、TPR(θ) および FPR(θ) は以下のように計算される。FPR が低い状態で TPR が高いモデルが良いモデルである。

$$\text{TPR}(\theta) = \frac{\text{TP}(\theta)}{\text{TP}(\theta) + \text{FN}(\theta)} \quad (4)$$

$$\text{FPR}(\theta) = \frac{\text{FP}(\theta)}{\text{FP}(\theta) + \text{TN}(\theta)} \quad (5)$$

AUC は曲線下の面積を表し、0 以上 1 以下の範囲で数値が高いほど精度が良い。なお、スコアリングによる並び替えではなく、一様ランダムでシャッフルした順序を用いた場合の AUC は 0.5 となる。

5.2 予測手法

上述したタスクに対して、複数の予測モデルの比較評価を行う。本節では予測モデルの詳細について説明する。

(1) **人気スコア法**: 最も単純なベースラインとして、全ユー

表 3: TP, TN, FP, FN の定義

		モデルの予測	
		適合	不適合
正解	適合	TP(θ) (真陽性)	FN(θ) (偽陰性)
	不適合	FP(θ) (偽陽性)	TN(θ) (真陰性)

ザに対して人気スコアに基づく同じ推薦を行うと仮定する。具体的には、各展示物の動画視聴時間の割合の合計値を計算し、0~1 で min-max 正規化した値を人気スコアとした。なお値が大きいほど人気が高いとみなせる。閾値を変化させた場合の動画視聴履歴の予測精度を評価する。

(2) **動画視聴有無の履歴ベース法**: 2.3 節で述べた NMF で近似分解したアイテム埋込み行列 \mathbf{V} を用いて、動画視聴有無の履歴を予測する。まず、式 (3) と類似の方法で、ユーザ \times 展示物のペアの動画視聴有無を表す行列 \mathbf{R} を作成する。行列 \mathbf{R} において $m = \sum_{y=1}^{n-1} |U_y|$ であり、予測対象となるユーザ以外で、展示物の動画を x 回以上視聴した履歴のあるユーザのデータを用いる。行列 \mathbf{R} の要素は動画視聴有無を表し、ユーザ i が展示物 j の動画を視聴した場合は $r_{ij} = 1$ 、視聴しなかった場合は $r_{ij} = 0$ とする。なお、動画を視聴したかどうかの判定には、動画視聴時間の割合が w より大きいかどうかで決定した。ここで、予測対象となるユーザ z の動画視聴有無の予測結果を \mathbf{x}_z^* 、 x 箇所動画を視聴した現在の履歴を $\mathbf{u}_z = (u_{z1}, u_{z2}, \dots, u_{z|E|})^\top$ とする。 \mathbf{u}_z の要素は、動画を視聴した場合に 1、それ以外の場合に 0 となる。次に、行列 \mathbf{R} を式 (1) に従って近似分解し、アイテム埋込み行列 \mathbf{V} を得る。 \mathbf{x}_z^* は \mathbf{V} とユーザ z の埋込み行列の内積で求めることができ、ユーザ z の埋込み行列は $(\mathbf{V}\mathbf{V}^\top)^{-1}\mathbf{V}\mathbf{u}_z$ で算出する。よって、動画視聴有無の予測スコア \mathbf{x}_z^* は以下のような式で求める。

$$\mathbf{x}_z^* = \mathbf{V}^\top (\mathbf{V}\mathbf{V}^\top)^{-1} \mathbf{V}\mathbf{u}_z \quad (6)$$

なお NMF における基底数 k は 10 とし、 $x \in \{1, 2, 3\}$ 、 $w \in \{0.0, 0.3, 0.5\}$ とした。

(3) **動画視聴割合の履歴ベース法**: 動画視聴有無の履歴ベース法と同様の方法で、動画視聴有無の履歴を予測する。ただし、行列 \mathbf{R} はユーザ \times 展示物のペアの動画視聴時間の割合を表し、行列の各要素は $r_{ij} \in \mathbb{R}_+$ で表現する⁵。予測対象となるユーザ z の動画視聴有無の予測結果を \mathbf{x}_z^* とし、 x 箇所動画を視聴した現在の履歴を $\mathbf{u}_z = (u_{z1}, u_{z2}, \dots, u_{z|E|})^\top$ とする。なお、 \mathbf{u}_z の要素は動画視聴時間の割合を表し、未視聴の動画に対応する展示物の要素は 0 とする。動画視聴有無の予測スコア \mathbf{x}_z^* は式 (6) で述べた式で求める。NMF における基底数 k は 10 とし、 $x \in \{1, 2, 3\}$ とした。

別日の同じ ID を持つユーザは別ユーザとして扱い、 $x+1$ 以上の動画を視聴したユーザを予測対象とした。予想対象となるユーザ数を図 9 に示す。なお、動画視聴割合の履歴ベース法に

⁵: \mathbb{R}_+ は 0 を含む正の実数を表す。

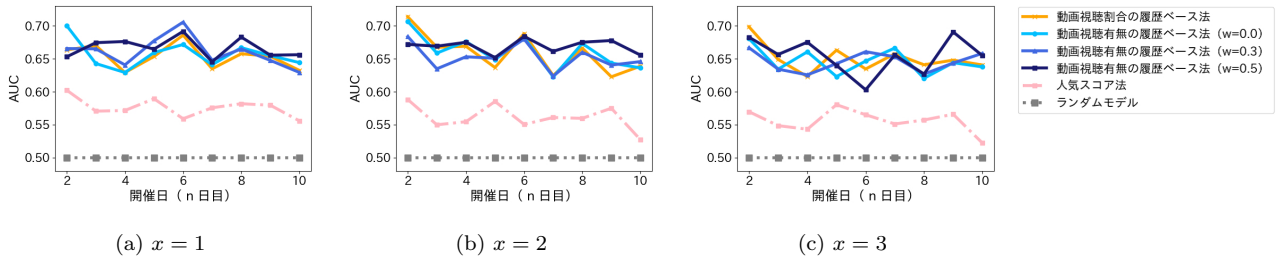


図 8: 動画視聴有無の履歴を予測した結果の AUC の比較 ($w > 0.5$ で動画を視聴したと判定)

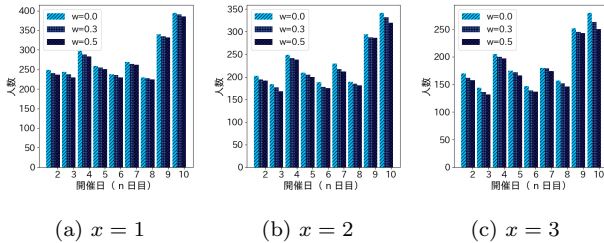


図 9: 各条件での予想対象となるユーザ数

における予測対象ユーザ数は $w = 0$, 人気スコア法は $w = 0.5$ の場合と同じである。

5.3 評価結果

各ユーザの AUC を日毎に平均した結果を図 8 に示す。横軸は開催日 ($2 \leq n \leq 10$) を表し、縦軸は AUC を表す。動画視聴割合および視聴有無の履歴ベース法は、いずれもランダムモデルおよび人気スコア法よりも高い精度となった。よって、ユーザの動画視聴履歴を用いた展示物の推薦が効果的であることが示された。予測に用いる動画視聴数 x の値を増やすと精度が向上すると期待されたが、いずれのモデルでも x の値は精度向上に寄与しなかった。特に人気スコア法に着目すると、ゆるやかに精度が低下している。これは、後半に視聴する動画ほど興味の度合いが小さくなり、履歴を予測するのが難しくなるためだと考えられる。また、動画視聴有無を判定するパラメータ w に着目しても精度の違いは見られず、動画視聴時間の割合の履歴ベース法も、これらのモデルと同様の精度であった。この結果により、興味の度合いを測るのに動画の視聴時間はほぼ関係ないと言える。これは動画を視聴開始した時点で十分に興味があると言えることを意味するが、この傾向はオンラインカンファレンス特有のものだと考えられる。また、日付が進むにつれて予測に用いるユーザデータが増加するため精度が向上すると予想されたが、日付による精度の違いは見られなかった。この要因として、予測に用いたユーザデータに日々偏りがあったことが考えられる。開催日初日のデータのみを用いて残りの開催期間を予測したとしても、十分に効果的な推薦が可能だと言える。

5.4 まとめ

RQ3 (有効性) に関するまとめを述べる。取得したユーザのログデータを用いたオフライン評価により、NMF を用いた

手法はランダムモデルおよび人気スコア法よりも高い精度で展示物の動画視聴履歴を推定できることが示された。このことから、情報推薦により、オンラインカンファレンスにおいてユーザの興味に適合した展示物を提示でき、適合性の観点でユーザの満足度の向上が期待できる。

オンラインカンファレンスにおける実際の情報推薦のタスクは履歴を推定することではなく、ユーザの顕在的および潜在的な興味に適合した展示物を予測することであるため、本稿で想定したタスクとは差分がある。特に潜在的な興味に関しては、今回実施したオフライン評価では考慮できていない。そのため今後はオンライン評価を実施し、セレンディピティに着目した検証が必要である。

6 おわりに

本研究ではオンラインカンファレンスにおいて情報推薦が有効であるかを議論した。取得されたオンラインカンファレンスのログデータを用いてユーザの行動を分析した。また NMF を用いた手法についてオフライン評価を実施した。その結果、展示物の動画視聴履歴をランダムモデルおよび人気スコア法よりも高い精度で推定できた。これらのことから、オンラインカンファレンスにおいて情報推薦によりユーザの興味に適合した展示物を提示できることが示され、適合性の観点でユーザの満足度向上が期待できる。一方で潜在的な興味に関しては今回実施したオフライン評価では考慮できていないため、オンライン評価によるセレンディピティに着目した検証が今後の課題である。

文献

- [1] Robin Burke. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370, 2002.
- [2] Folasade Olubusola Isinkaye, YO Folajimi, and Bolande Adewoke Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian informatics journal*, 16(3):261–273, 2015.
- [3] Vivek Kalia, Ashok Srinivasan, Luke Wilkins, and Gary D Luker. Adapting scientific conferences to the realities imposed by covid-19, 2020.
- [4] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.
- [5] Denis Kotkov, Jari Veijalainen, and Shuaiqiang Wang. Challenges of serendipity in recommender systems. In *International conference on web information systems and tech-*

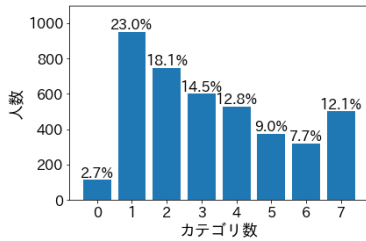


図 10: 訪問カテゴリ数

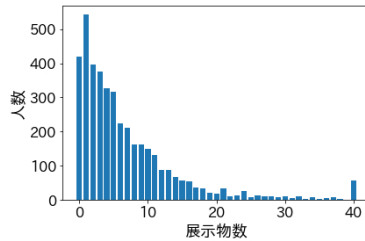


図 11: 訪問テーマ数

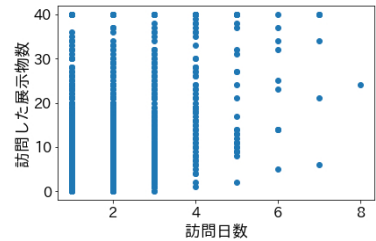


図 12: 訪問テーマ数と訪問日数

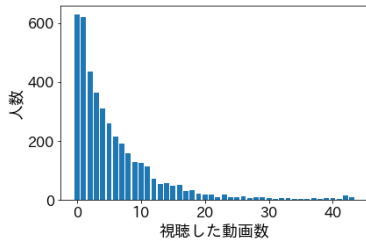


図 13: 動画視聴数

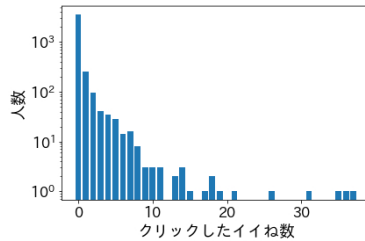


図 14: イイねクリック数

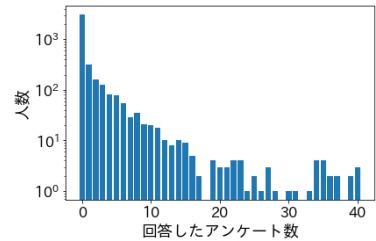


図 15: アンケート回答数

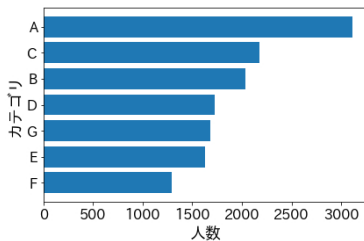


図 16: 各カテゴリの訪問数の比較

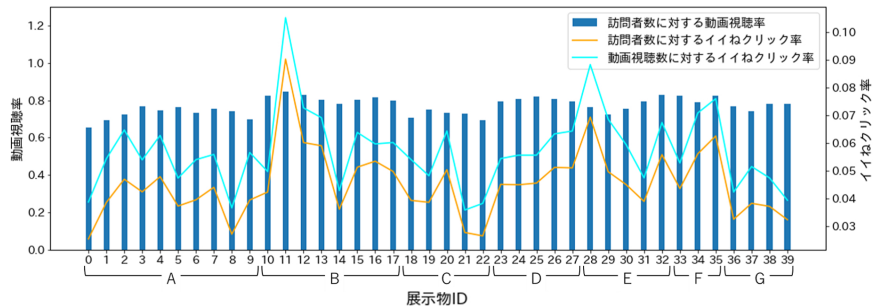


図 17: 各展示物ページの訪問者数に対する動画視聴率とイイね率、および動画視聴数に対するイイね率

nologies. SCITEPRESS, 2016.

- [6] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.
- [7] Takeaki Uno, Masashi Kiyomi, and Hiroki Arimura. Lcm ver. 3: Collaboration of array, bitmap and prefix tree for frequent itemset mining. In *Proceedings of the 1st international workshop on open source data mining: frequent pattern mining implementations*, pages 77–86, 2005.
- [8] Antonio Valenti, Grazia Fortuna, Caterina Barillari, Erika Cannone, Valeria Boccuni, and Sergio Iavicoli. The future of scientific conferences in the era of the covid-19 pandemic: Critical analysis and future perspectives. *Industrial health*, 2021.
- [9] 奥健太. セレンディピティ指向情報推薦の研究動向. *知能と情報*, 25(1):2–10, 2013.
- [10] 神島敏弘. 推薦システムのアルゴリズム. <https://www.kamishima.net/archive/recsysdoc.pdf>, アクセス日: 2021.12.29.

付 録

取得されたオンラインカンファレンスにおけるユーザのログデータについて、ユーザの行動および展示物に関する基本的な統計情報を付録にまとめる。

a) ユーザの行動に関する統計情報

図 10 および図 11 は、各ユーザが訪問したカテゴリ数および展示物数を表す。訪問したカテゴリ数には、展示物ページおよびカテゴリページの両方が含まれる。図 12 はユーザが web サイトを訪問した日数と訪問した展示物数の散布図を示す。図 13 は各ユーザの動画視聴数、図 14 はイイねボタンのクリック数、図 15 は展示物へのアンケート回答数を表す。図 14 および図 15 の縦軸は対数である。なお一部のカテゴリページのみ動画があるため、カテゴリページ数と展示物ページ数の合計と動画数は一致しない。

b) 展示物に関する統計情報

図 16 はユーザの訪問数が多かったカテゴリのランキングを表す。横軸は各カテゴリページおよびそこに含まれる展示物ページに訪問した人数を表している。図 17 は、各展示物ページの訪問者数に対する動画視聴率とイイねクリック率、および動画視聴数に対するイイねクリック率を表す。左縦軸は動画視聴率を表し、右縦軸はイイねクリック率を表す。