

Synapse: 利用者の文脈に応じて 継続的に推薦手法の選択を最適化する推薦システム

三宅 悠介^{1,a)} 松本 亮介^{2,b)}

概要: EC サイトの商品種類増大に伴う情報過多問題を解決するため、利用者の要求を満たす商品を自動的に提案する機能が EC サイトに導入されている。この機能に用いられる推薦手法は数多く提案されていることから、EC サイトの運営者にとって利用者の要求を満たす可能性の高い推薦手法を選択することが重要となる。しかしながら、推薦手法の優劣は扱う商品の特性を含む多くの文脈ごとに変化するため、実環境での評価なしに推薦手法を決定することは難しい。また、商品を自動的に提案する機能の効果を最大化するため、新しい推薦手法や優劣を助長する文脈を導入し継続的に評価することが求められる。本報告では、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する推薦システムを提案する。提案手法では、推薦手法の選択の最適化を多腕バンディット問題とみなし Epsilon-Greedy アルゴリズムを用いて解く。この最適化を、定めた文脈ごとに定期的に施すことで新しい推薦手法や文脈の導入に継続的に適応する。提案手法の評価のため、EC サイトにおいて閲覧中の商品カテゴリごとに有効な推薦手法が異なることを確認した。次に閲覧中の商品カテゴリを文脈と見なした最適化によって、文脈を考慮しない場合と比較して累積クリック数が向上することを確認した。

Synapse: Recommendation system to continuously optimize selection of recommendation method based on user context

YUSUKE MIYAKE^{1,a)} RYOSUKE MATSUMOTO^{2,b)}

Abstract: A recommendation system is introduced to solve the information overload problem of the electronic commerce sites. EC site should select efficiency one from many recommendation methods being proposed. Because effective method is different based on various contexts, deciding the suitable method without evaluation on the site is difficult. In addition, EC sites should continuously evaluate new methods and contexts to improve the system. In this report, we propose recommendation system to continuously optimize selection of recommendation method based on context defined in advance. We consider optimization selection of method as multi-armed bandit problem and resolve the one using Epsilon-Greedy algorithm. The proposed system adapts continuously by regularly performing this optimization. For evaluation, we measured that the effective method for each product categories on a EC site was different. Next, we measured that the cumulative click count improved compared to the one without context.

1. はじめに

EC サイト市場規模は成長を続けており [20], それに伴い取り扱う商品の種類は増大している。EC サイト利用者の通常の行動では全ての商品を見て回ることは困難であるため、多くの EC サイトでは膨大な商品の中から利用者が興味を持つであろう商品を自動的に提案する機能を導入している。この機能では、利用者に提案する商品を選定する

¹ GMO ペパボ株式会社 ペパボ研究所
Pepabo R&D Institute, GMO Pepabo, Inc., Tenjin, Chuo-ku, Fukuoka 810-0001 Japan

² さくらインターネット株式会社 さくらインターネット研究所
SAKURA Research Center, SAKURA Internet Inc., Akasaka, Chuo-ku, Fukuoka 810-0042 Japan

a) miyakey@pepabo.com

b) r-matsumoto@sakura.ad.jp

ためのアプローチを決め、採用する。これらのアプローチは推薦手法と呼ばれ、多くの手法が提案、体系化されてきた [3], [5], [14], [15]。そこで、EC サイトの運営者にとっては、数多く提案されている推薦手法の中から利用者の要求を満たす可能性の高い手法を選択することが重要となる。

効果的な推薦手法の選択のためには、文脈ごとの推薦手法の使い分けと継続的な評価が重要である。ここでの文脈とは、利用者を含む推薦システムの状態を特徴付ける様々な情報のうち、推薦手法の選択に影響を及ぼす情報を言う。そのような文脈として、推薦手法の情報源やその状態が挙げられる。推薦手法は、推薦のための情報源として、利用者の嗜好や人口統計学的情報、商品の属性情報や嗜好情報と属性情報の関連性を記述した知識モデル、利用者や商品同士の相関情報などを用いる。しかしながら対象となる利用者や商品に応じた情報源の整備または蓄積具合の差によって、同じ推薦手法においても推薦結果が異なる。そのため、推薦手法の比較では、それぞれが用いる情報源の状態によって相対的な優劣が変化し得る [5]。また、我々は、商品画像の類似画像による推薦手法と利用者同士の嗜好の類似性を用いた推薦手法の比較において、商品の特性によって推薦手法の優劣が異なることを確認した [19]。このことから、推薦対象となる商品の特性も同様に文脈とみなすことができる。利用者の要求を満たす可能性の高い推薦手法を選択するためには、これらの情報源の状態や扱う商品の特性を始めとする様々な文脈を見定め、これに適した推薦手法を選択することが重要である。ただし、これらの文脈は、推薦手法の評価を行う環境に依存することから、実環境での評価なしに決定することは難しい。

実環境での評価では、評価期間中の機会損失が課題となる。このような機会損失が発生する状況として、評価の劣る推薦手法を候補として利用し続ける場合や、評価の見切りが早過ぎることで長期的な評価で勝る推薦手法を利用しない場合がある。これらの機会損失を抑えるために、ある時点での評価の高い推薦手法を利用しながら、並行して限定的に他の候補との評価を行う必要がある。加えて、EC サイトでは、商品を自動的に提案する機能の効果を最大化するため、新しい推薦手法や文脈の導入に対しての継続的な評価が求められる。そのため、効果的な推薦手法の選択のためには、多くの文脈と推薦手法の組み合わせに対する機会損失の抑えた継続的な評価が望ましい。

本研究では、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する推薦システムを提案する。提案手法では、推薦手法の選択に影響を及ぼす文脈を仮定し、その文脈ごとに推薦手法の選択を行う。文脈ごとの最適な推薦手法は実環境での評価によって求めるが、機会損失を抑えるために評価の高い推薦手法の利用と他の候補の評価を並行する。この利用と評価のトレードオフの最適な解を求めるために、同評価を多腕バンディット問題 [10] と

みなし、この解法の一つである Epsilon-Greedy アルゴリズム [16] によって解く。これらの一連の評価を定期的に実行することで、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する。

本論文の構成を述べる。2 章では、利用者の要求を満たす可能性の高い推薦手法の選択に関する課題について述べる。3 章では、2 章で述べた文脈ごとの推薦手法の最適化の課題を解決するための提案システムの実装を述べる。4 章では、文脈ごとの推薦手法の選択の最適化手法の評価を行い、5 章でまとめとする。

2. 推薦手法の選択

推薦手法は数多く提案されていることから、EC サイトの運営者にとって利用者の要求を満たす可能性の高い推薦手法を選択することが重要となる。本章では EC サイトにおける推薦手法の選択に関する課題について整理する。

2.1 推薦手法の選択と文脈

利用者の要求を満たす可能性の高い推薦手法を選択するためには、文脈ごとの推薦手法の使い分けが重要となる。本報告では、文脈をコンテキストウェア推薦における定義 [1] のそれを、本報告の目的に照らして限定する。すなわち、利用者を含む推薦システムの状態を特徴付ける情報 [13] のうち、推薦手法の選択に影響を及ぼす情報を文脈と呼ぶ。

このような文脈として、推薦手法の情報源やその状態を用いることができる。EC サイトで用いられる代表的な推薦手法の方式として、商品の属性情報を用いる内容ベース推薦 [12]、嗜好情報と属性情報の関連性を記述した知識ベース推薦 [6]、利用者や商品同士の相関情報を用いる協調型推薦 [8] が挙げられる。これらは、対象となる利用者や商品に応じた情報源やその整備または蓄積具合の差によって、同じ推薦手法においても推薦結果が異なる。例えば、内容ベース推薦や知識ベース推薦では、推薦結果が情報源との直接の類似性によって得られることから、推薦内容の意外性が低下することが知られている [3], [5], [12]。また、協調型推薦では、情報源となる嗜好情報が不足する状況で推薦精度の低下や新アイテムの推薦対象に含まれない、いわゆるコールドスタート問題 [9] が存在する。このため、推薦手法の比較では、それぞれが用いる情報源とその状態によって相対的な優劣が変化し得る。推薦手法を組み合わせることで各手法の短所を補う手法であるハイブリッド型推薦 [5] は、この文脈に着目した推薦手法の使い分けと言える。

同様に、推薦対象となる商品の特性も文脈とみなすことができる。我々は、商品画像の類似画像を用いた内容ベース推薦と利用者同士の嗜好の類似性を用いた協調型推薦の比較において、商品の特性によって推薦手法の優劣が異な

ることを確認した [19]。この報告では、あみぐるみ、セーター・カーディガンを扱う商品カテゴリにおいて、類似画像を用いた内容ベース推薦が、形状や質感の特徴を見出すことで既存の商品属性情報を補完し、利用者にとって効果的な商品が選定することがわかった。一方、しおり・ステッカーやコスチューム、イラストといった、形状やデザインに統一性がなく商品画像の観点からの類似性が低い商品カテゴリでは、類似画像を用いた内容ベース推薦と比べて、商品の属性情報に依存しない利用者同士の嗜好の類似性を用いた協調型推薦が良い結果を得た。

利用者の要求を満たす可能性の高い推薦手法を選択するためには、これらの情報源の状態や扱う商品の特性を始めとする様々な文脈を見定め、これに適した推薦手法を選択することが重要である。

2.2 推薦手法の選択の基準

推薦手法の選択に影響を及ぼす文脈を見出し、これに適した推薦手法を選択するには判断基準が必要である。推薦手法の情報源とその状態を文脈とみなし、有効な推薦手法を使い分けるハイブリッド推薦では、切り替えのために推薦手法の精度 [5] や利用者の評価 [17] を用いる。これらの具体的な値は利用者を含む環境によって異なることから、推薦手法を導入する実環境における推薦手法の評価が求められる。商品の特性を文脈とみなす場合も、EC サイトによって取り扱う商品の特性が異なることから同様に実環境での評価が必要である。

推薦行為が利用者の要求を満たしたことを判断する指標として、推薦結果に対して利用者が閲覧もしくは購入した割合を計測したものが用いられる。これらはそれぞれクリック率 (CTR) やコンバージョン率 (CVR) と呼称される。文脈に適した推薦手法を選択する明確な判断基準を得るためには、実環境における指標の実際の計測が必要となる。

2.3 推薦手法の継続的な選択

EC サイトでは、商品を自動的に提案する機能の効果を最大化するため、新しい推薦手法や文脈を導入し継続的に評価を行う必要がある。評価では、新しい推薦手法や文脈が実環境に導入され、一定期間、指標となるクリック率やコンバージョン率が計測される。指標の計測は、評価時期の差異や評価者の偏りが発生しないよう、同期間かつ無作為に評価者を分け、計測対象の回数が均等になるようにして行われる。このような計測の方式は A/B テストと呼ばれている [11]。次に、計測で得られた指標の結果に基づき相対的に成績が高かった候補を選択し、これを用いるようシステムの振る舞いを変更する。

候補となる推薦手法の間に指標に対する優劣がある場合、成績の低い推薦手法を計測する期間が長くなることは、

成績の高い推薦手法を用いた時に対して機会損失が発生する。そのため、実環境での評価においては機会損失を抑えるために、この一連の工程は速やかに行う必要がある。文脈に応じた評価が必要な環境においては、これを考慮しない環境と比べ評価対象が増えることからこの工程の自動化が望ましい。また、新しい推薦手法や文脈の導入以外でも、時刻や季節によって推薦手法の優劣が変化する状況では、ある時点での成績による選択は長期的な観点での機会損失につながる恐れがある。そのため、この工程は継続的に実行される必要がある。

3. 提案手法

推薦手法の優劣が文脈によって変化し、その明確な基準が実環境での評価によって明らかになる環境において、継続的に追加される推薦手法や文脈を交えながら、利用者の要求を満たす商品を選定する可能性の高い推薦手法を選択する推薦システムにするためには、以下の要件が必要である。

- (1) 文脈を定めて、定めた文脈ごとに推薦手法を評価する
- (2) 文脈ごとの評価結果に基づき、推薦手法を使い分ける
- (3) これらの評価と使い分けを継続的に行う

本研究では、上述した 3 つの要件を満たすために、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する推薦システムを提案する。図 1 に提案システムの処理フローを示す。提案システムは、文脈ごとに選択可能な推薦手法から選定された推薦結果に対する利用者の行動を、文脈ごとの推薦手法に対する評価として記録する。次に、提案システムは蓄積された評価から、ある文脈において、その時点で最も実環境での評価の高い推薦手法を用いるよう振る舞いを変更する。ただし、推薦手法の追加や、時間経過による推薦手法の優劣が変化する可能性を考慮して、評価の低い推薦手法も並行して限定的に用いる。提案システムでは、この推薦手法の選択の最適化を多腕バンディット問題とみなし、この基本的な解法である Epsilon-Greedy アルゴリズムによってこれを解く。これらの一連の評価と結果反映が定期的に実行されることで、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する。

3.1 文脈と推薦手法の定義

提案手法では、文脈を、利用者を含む推薦システムの状況の特徴付ける様々な情報のうち、推薦手法の選択に影響を及ぼす情報と定義する。そして、推薦手法の優劣を助長するであろう文脈があることを仮定する。例えば、推薦対象の商品特性によって推薦手法の優劣が異なる場合、その特性をグループ化した商品カテゴリを文脈とみなすことができる。そこで、提案手法では、推薦システムで表現可能な文脈の中から、対象とする各文脈において選択可能な推

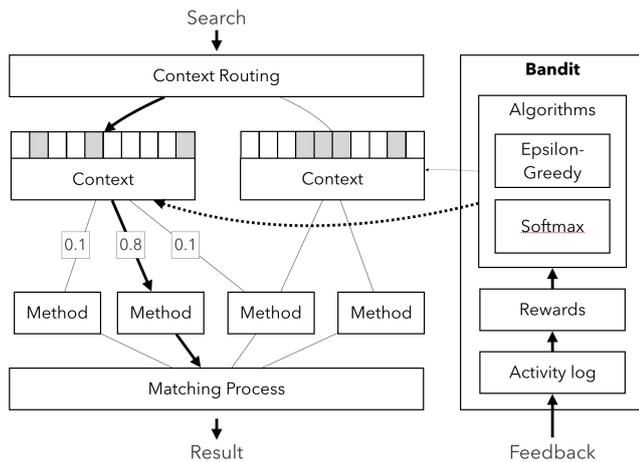


図 1 Synapse の処理フロー
Fig. 1 Flow of Synapse.

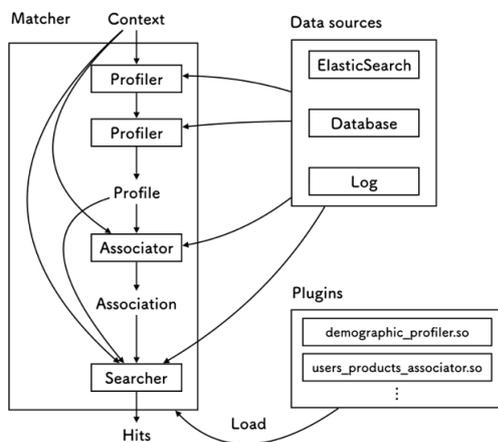


図 2 推薦手法のモジュール構成
Fig. 2 Recommendation method as Synapse module.

薦手法の一覧を定義しておく。利用者からの推薦要求に対しては、該当する文脈を判断し、これに紐づく候補の推薦手法から1つを選び、この推薦手法から得られた推薦結果を返す。

また、提案システムでは、複数の推薦手法を扱う必要から、導入時のアルゴリズム詳細に起因する個別対応を避けるため、共通のインターフェースを持つモジュールとして推薦手法を実装する。同様に、複数の推薦手法のモジュール開発を効率化するため、利用者の情報取得や関連付け、検索処理といった推薦に必要な処理を共通利用可能なフィルタとして定義し、これを組み合わせて推薦手法モジュールを開発する。推薦手法のモジュール構成を図2に示す。

3.2 文脈ごとの推薦手法の評価と使い分け

提案システムでは、推薦手法に対する評価を行うために推薦に対する利用者の反応を記録する。そのために提案システムは、利用者の任意の商品に対する閲覧や購入といっ

た行動時に、時刻、行動種別、対象の商品を記録する。また、利用者からの推薦要求に対しても、時刻、文脈の種類、選択された推薦手法、ならびにその推薦結果である商品一覧を記録する。推薦結果に対する利用者の反応は、利用者の推薦要求ごとに、その時刻以降で直近の閲覧や購入行動における商品が、推薦結果の商品一覧に含まれたかどうかをもって判断する。ここで直近の行動の行動種別を限定することで、閲覧であればクリック率、購入であればコンバージョン率として評価に用いることができる。

提案手法では、文脈ごとの推薦手法の選定を継続的に最適化させる必要から、評価の高い推薦手法の利用と候補の推薦手法の評価を並行して行う。評価は実環境で行われることから、機会損失を抑えるために、利用と評価のトレードオフを考慮しなければならない。多腕バンディット問題は、腕と呼ばれる複数の候補から得られる報酬を最大化する問題である。この問題の解法では、ある時点の腕ごとの評価に基づき、最も評価の高い腕を用いながらも、真に評価の高い腕の探索を並行して行う。ここで、提案手法における推薦手法の選定の継続的な最適化は、多腕バンディット問題とみなすことができる。提案手法では、文脈ごとの推薦手法の継続的な最適化のため、同問題の基本的な解法である Epsilon-Greedy アルゴリズムを用いる。多腕バンディット問題に対する解法では、ある時点で最も評価の高い腕を用いることを活用、各腕の評価を行うことを探索と呼ぶ。Epsilon-Greedy アルゴリズムでは、探索を行う割合を $0 \leq \epsilon \leq 1$ で表現し、探索時には候補を均等に出し分ける。

提案システムは、推薦手法の評価とその評価に基づく推薦手法の選定の最適化を定期的に行う。これにより、文脈ごとに推薦手法の選択を自動的かつ継続的に最適化が可能となる。

4. 評価と考察

本手法の有効性を確認するために、文脈ごとに推薦手法の優劣が確認できた環境から取得したデータを用いて推薦手法の選択の最適化処理を評価した。表1に評価に用いたデータを示す。本データはあるECサイトの推薦システムで利用されている5つの推薦手法とクリック率を表している。クリック率は、推薦手法が用いられた17の商品カテゴリごとに計測されており、それぞれのカテゴリにおいてクリック率や有効な推薦手法が異なっている。本評価では、この推薦手法が用いられた17の商品カテゴリを提案手法における文脈とした。なお、各商品カテゴリの推薦手法のうち最も高いクリック率を太字で示している。

本評価では、乱数を用いたシミュレーション手法として知られるモンテカルロ法 [2] を用い、各推薦手法がクリック率を確率として報酬を獲得する場合に一定の試行回数を経過した後の予測クリック数である累積報酬額を比較する。

比較は、①推薦手法の選定の最適化を行わない、②文脈を考慮せず推薦手法の選定の最適化を行う、そして提案手法である③文脈を考慮して推薦手法の選定の最適化を行う、の3つのシミュレーションに対して行った。①推薦手法の選定の最適化を行わないシミュレーションでは、5つの推薦手法を均等に用いる。これは Epsilon-Greedy アルゴリズムにおける探索率 ϵ が 1.0 の状態と同等である。②文脈を考慮せず推薦手法の選定の最適化を行うシミュレーションでは、表 1 の All 列にある全体のクリック率を確率として用いる。本シミュレーションでは全ての文脈において平均的に有効な推薦手法が活用される。提案手法である③文脈を考慮して推薦手法の選定の最適化を行うシミュレーションでは、文脈ごとのクリック率を確率として用いる。本シミュレーションでは各文脈における有効な推薦手法が活用される。また、本シミュレーションでは全体の試行回数がそれぞれの文脈で均等に分割されるものとし、試行ごとに決まった順序で文脈を切り替えて評価した。推薦手法の選択の最適化に用いた Epsilon-Greedy アルゴリズムの探索率である ϵ は 0.2、試行回数は 85,000 回とした。文脈数が 17 であることから、③文脈を考慮して推薦手法の選定の最適化を行う場合は各文脈において試行回数が 5,000 回となる。なお、乱数を用いた確率の計算結果を平均化するために上述のシミュレーションを 500 回行い、この平均を結果として用いた。

図 3 に試行回数と累積報酬額のシミュレーション間の比較を示す。①推薦手法の選定の最適化を行わない場合の累積報酬額は 8,438 回、②文脈を考慮せず推薦手法の選定の最適化を行う場合は 11,623 回、③文脈を考慮して推薦手法の選定の最適化を行う場合は 12,396 回であった。Epsilon-Greedy アルゴリズムを用いた推薦手法の選定の最適化によって、最適化を行わない場合と比べて累積報酬額であるクリック総数が増加することが確認できる。また、提案手法である文脈ごとの推薦手法の選定の最適化によって、文脈を考慮しない場合と比べてクリック数が増加することがわかる。

次に、同シミュレーションにおける試行回数のうち開始 1,000 回の累積報酬額を切り出したものを図 4 に示す。②文脈を考慮せず推薦手法の選定の最適化を行う場合の累積報酬額は 126 回に対して、③文脈を考慮して推薦手法の選定の最適化を行う場合は 114 回であり、提案手法による最適化による累積報酬額が低い期間があることがわかる。これは、提案手法が文脈ごとに最適化を行うことから文脈単位での試行回数が必然的に少なくなるためと考えられる。

最後に、シミュレーションに用いるクリック率が低い状態での累積報酬額を図 5 に示す。この状態を再現するために表 1 に示すクリック率に一律 0.1 を乗じた。また、十分な累積報酬額へ到達させるため試行回数を 3 倍としている。②文脈を考慮せず推薦手法の選定の最適化を行う場合

表 1 推薦手法のクリック率

Table 1 CTR of recommendation methods.

Category	C	D	L	I	T
1	0.10	0.02	0.09	0.12	0.14
10	0.17	0.01	0.09	0.16	0.16
21	0.13	0.01	0.08	0.15	0.16
27	0.09	0.01	0.10	0.13	0.15
32	0.09	0.01	0.10	0.13	0.18
39	0.25	0.01	0.12	0.19	0.14
45	0.06	0.01	0.09	0.13	0.12
50	0.13	0.01	0.07	0.11	0.15
58	0.05	0.00	0.08	0.12	0.13
64	0.20	0.01	0.13	0.15	0.24
69	0.18	0.01	0.11	0.17	0.14
74	0.13	0.00	0.14	0.14	0.16
79	0.05	0.02	0.08	0.12	0.18
84	0.10	0.01	0.09	0.16	0.15
99	0.09	0.01	0.09	0.13	0.14
110	0.03	0.01	0.10	0.12	0.08
163	0.00	0.01	0.09	0.12	0.14
All	0.12	0.01	0.09	0.13	0.15

の累積報酬額は 3,447 回に対して、③文脈を考慮して推薦手法の選定の最適化を行う場合は 3,565 回と、提案手法が累積報酬額で上回ったが、提案手法の累積報酬額が低い期間が前出のシミュレーションと比較して長くなっていることがわかる。これは、クリック率が一律に低く設定したことで、推薦手法間での差が縮まり、結果として個々の推薦手法の優劣の差が助長されない文脈となったためと考えられる。

本評価結果から、推薦手法の優劣を助長する文脈を適切に選定し、これを考慮した推薦手法の選定の最適化によって、考慮しない場合と比較して累積報酬額の向上に繋がることがわかった。一方で、提案手法では文脈ごとに最適化を行うことから、試行回数の少ない期間における累積報酬額に関する課題も確認した。この課題について、試行開始の時期には十分な探索回数を確保し、試行回数の増加に伴って探索回数を収束させるアニーリング処理 [18] や、既知の報酬を利用した効率的な探索を行う Softmax アルゴリズム [7] の採用を始めとする改善の検討が必要である。

5. まとめ

本研究では、事前に定めた文脈ごとに推薦手法の選択を自動的かつ継続的に最適化する推薦システムを提案し、その有効性を示した。実験から、推薦手法の選択に影響を及ぼす文脈を適切に選定し、これを考慮した推薦手法の選定の最適化によって、考慮しない場合と比較して累積報酬額の向上に繋がることがわかった。また、推薦手法の選定に必要な評価が文脈ごとに分散することから、評価の定まらない期間では結果が逆転する可能性が判明したことで、評

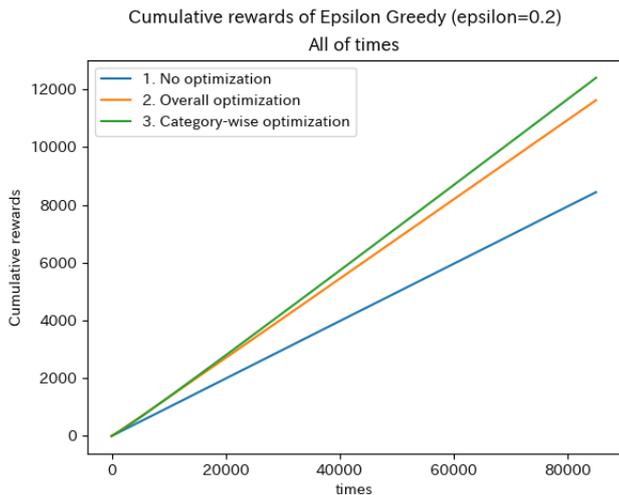


図 3 試行回数と累積報酬額のシミュレーション間比較
Fig. 3 Comparison of cumulative rewards for all times.

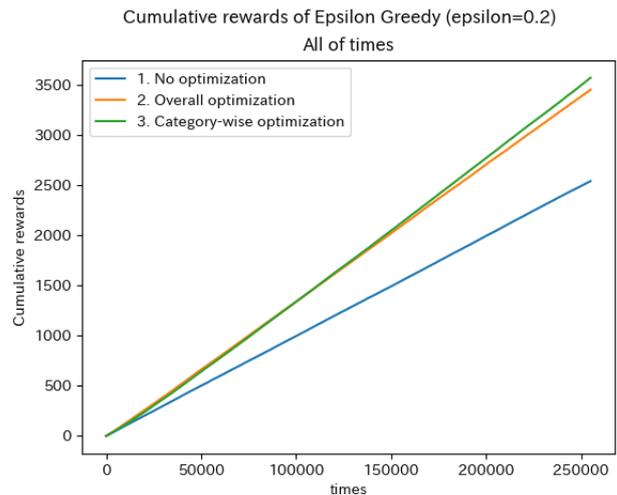


図 5 低確率時の試行回数と累積報酬額のシミュレーション間比較
Fig. 5 Comparison of cumulative rewards for all times in case of low probability.

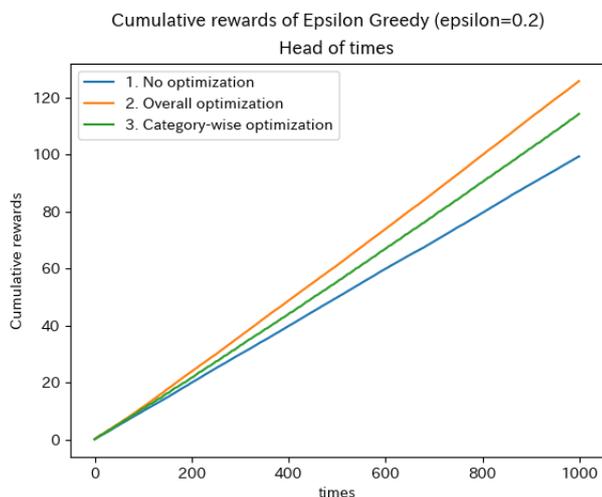


図 4 試行回数のうち開始 1,000 回と累積報酬額のシミュレーション間比較

Fig. 4 Comparison of cumulative rewards for 1,000 times.

価初期の探索率の調整が重要であることも示唆された。今後の課題として、文脈ごとの推薦手法の選定をプロダクション環境で行うことや、推薦手法の選択に影響を及ぼす効果の高い文脈の発見、ならびに特定の文脈において評価の高い推薦手法の確立による文脈ごとの効果向上の実現が挙げられる。

参考文献

[1] Abowd, Gregory D., et al. "Towards a better understanding of context and context-awareness." International symposium on handheld and ubiquitous computing. Springer, Berlin, Heidelberg, 1999.
[2] Binder, Kurt, et al. "Monte Carlo simulation in statistical physics." Computers in Physics 7.2 (1993): 156-157.
[3] Bobadilla, J., et al. "Recommender systems survey." Knowledge-based systems 46 (2013): 109-132.
[4] Broder, Andrei Z. "On the resemblance and contain-

ment of documents." Compression and Complexity of Sequences 1997. Proceedings. IEEE, 1997.
[5] Burke, Robin. "Hybrid recommender systems: Survey and experiments." User modeling and user-adapted interaction 12.4 (2002): 331-370.
[6] Burke, Robin. "Knowledge-based recommender systems." Encyclopedia of library and information science 69.Supplement 32 (2000): 180.
[7] Cesa-Bianchi, Nicolo, and Paul Fischer. "Finite-Time Regret Bounds for the Multiarmed Bandit Problem." ICML. Vol. 1998. 1998.
[8] Greg Linden, Brent Smith, Jeremy York, Amazon.com Recommendations: Item-to-Item Collaborative Filtering, IEEE Internet Computing Volume 7 Issue 1, pp. 76-80, January 2003
[9] Hyung Jun Ahn, A new similarity measure for collaborative filtering to alleviate the new user coldstarting problem, Information Sciences 178, pp. 37-51, 2008
[10] Katehakis, Michael N., and Arthur F. Veinott Jr. "The multi-armed bandit problem: decomposition and computation." Mathematics of Operations Research 12.2 (1987): 262-268.
[11] Kohavi, Ron, et al. "Controlled experiments on the web: survey and practical guide." Data mining and knowledge discovery 18.1 (2009): 140-181.
[12] Lops, Pasquale, Marco De Gemmis, and Giovanni Semeraro. "Content-based recommender systems: State of the art and trends." Recommender systems handbook. Springer, Boston, MA, 2011. 73-105.
[13] Ranganathan, Anand, and Roy H. Campbell. "An infrastructure for context-awareness based on first order logic." Personal and Ubiquitous Computing 7.6 (2003): 353-364.
[14] Schafer, J. Ben, Joseph Konstan, and John Riedl. "Recommender systems in e-commerce." Proceedings of the 1st ACM conference on Electronic commerce. ACM, 1999.
[15] Schafer, J. Ben, Joseph A. Konstan, and John Riedl. "E-commerce recommendation applications." Data mining and knowledge discovery 5.1-2 (2001): 115-153.
[16] Sutton, Richard S., and Andrew G. Barto. Introduction to reinforcement learning. Vol. 135. Cambridge: MIT

- press, 1998.
- [17] Tran, Thomas, and Robin Cohen. "Hybrid recommender systems for electronic commerce." Knowledge-Based Electronic Markets, Papers from the AAAI Workshop, Technical Report WS-00. Vol. 4. 2000.
 - [18] Vermorel, Joannes, and Mehryar Mohri. "Multi-armed bandit algorithms and empirical evaluation." European conference on machine learning, Springer, Berlin, Heidelberg, 2005.
 - [19] 三宅 悠介, 松本 亮介, 力武 健次, 栗林 健太郎, 特徴抽出器の学習と購買履歴を必要としない類似画像による関連商品検索システム, 研究報告インターネットと運用技術 (IOT) , Vol.2017-IOT-37(4), pp.1-8, May 2017
 - [20] 経済産業省 商務情報政策局 情報経済課, 平成 29 年度 我が国におけるデータ駆動型社会に係る基盤整備 (電子商取引に関する市場調査) , 2018