

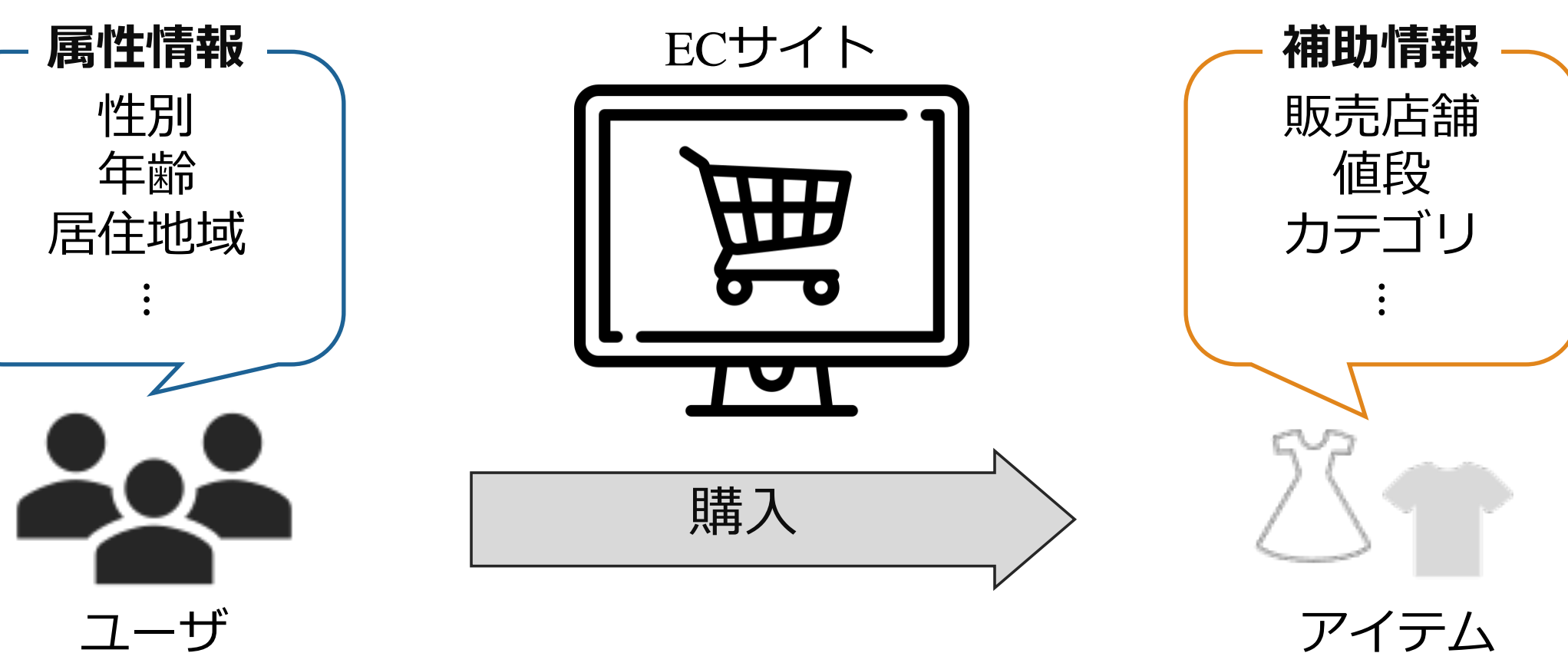
研究背景と目的

研究目的

ユーザの属性情報やアイテムの補助情報を統合的に扱い、顧客の購買行動を分析可能なモデルの構築および、構築したモデルを用いた分析方法の提案

現在、ECサイトではユーザの購買行動に関する

様々な情報を入手することが可能



提案への着想

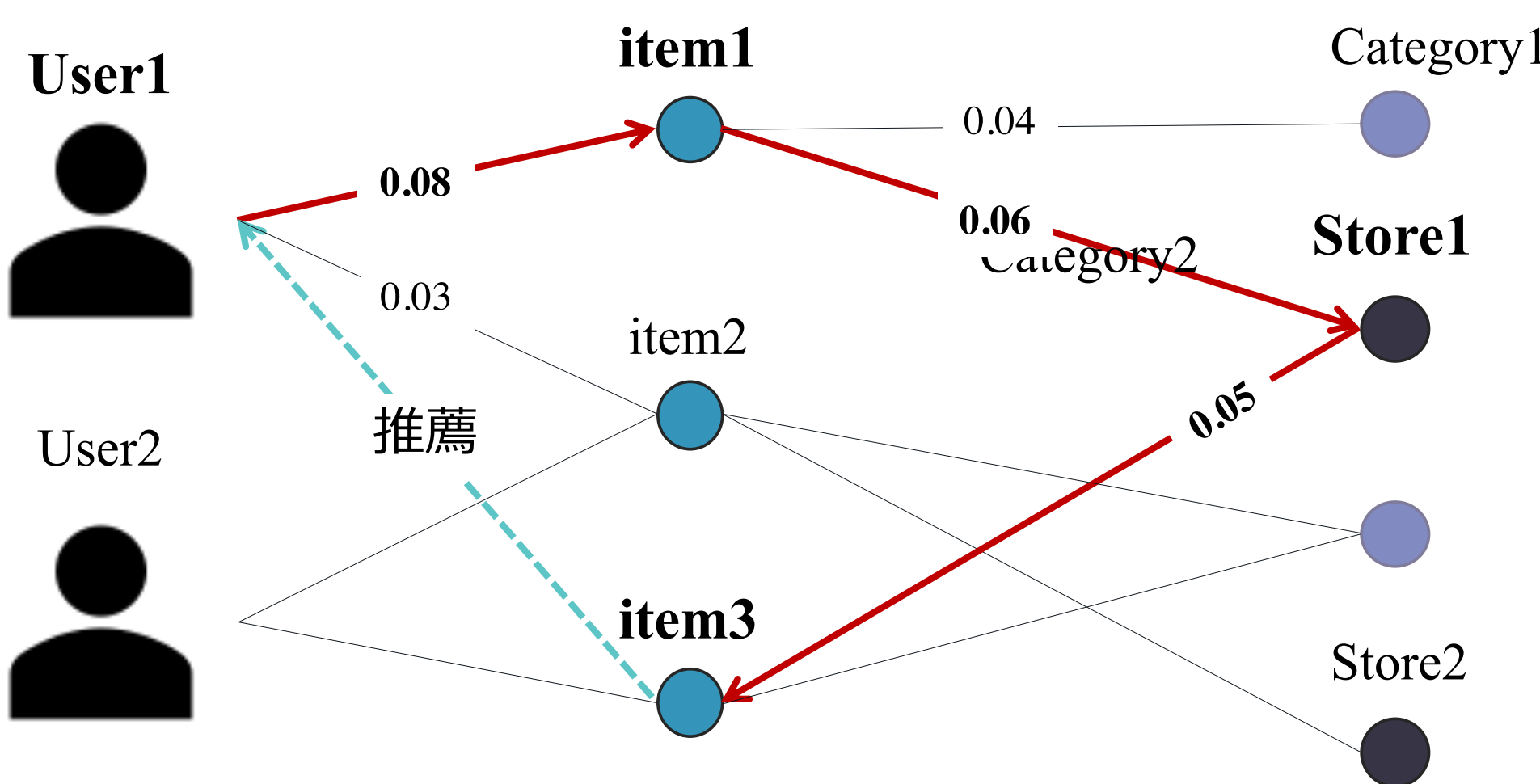
推薦システムの分野において、アイテムの補助情報を扱うことができ、解釈性を与えた手法として

Knowledge Graph Attention Network (以下、KGAT) が存在

- アイテムに関する補助情報を含んだグラフ構造のデータを入力とするモデル
グラフのエッジの重みを用いて、アイテムの推薦理由に対して解釈を行うことが可能

推薦モデルであるKGATの解釈性に着目し、顧客分析モデルとして再構築することを提案

例) User1にitem3が推薦されたのは、過去に購入したitem1と同じ店舗Store1で販売されているから



提案手法

KGAT をユーザの属性情報も取り扱うモデルへと拡張し、ユーザの属性情報・購買アイテム・アイテムの補助情報からユーザの購買行動を多角的に分析することが可能なモデルを提案

提案手法

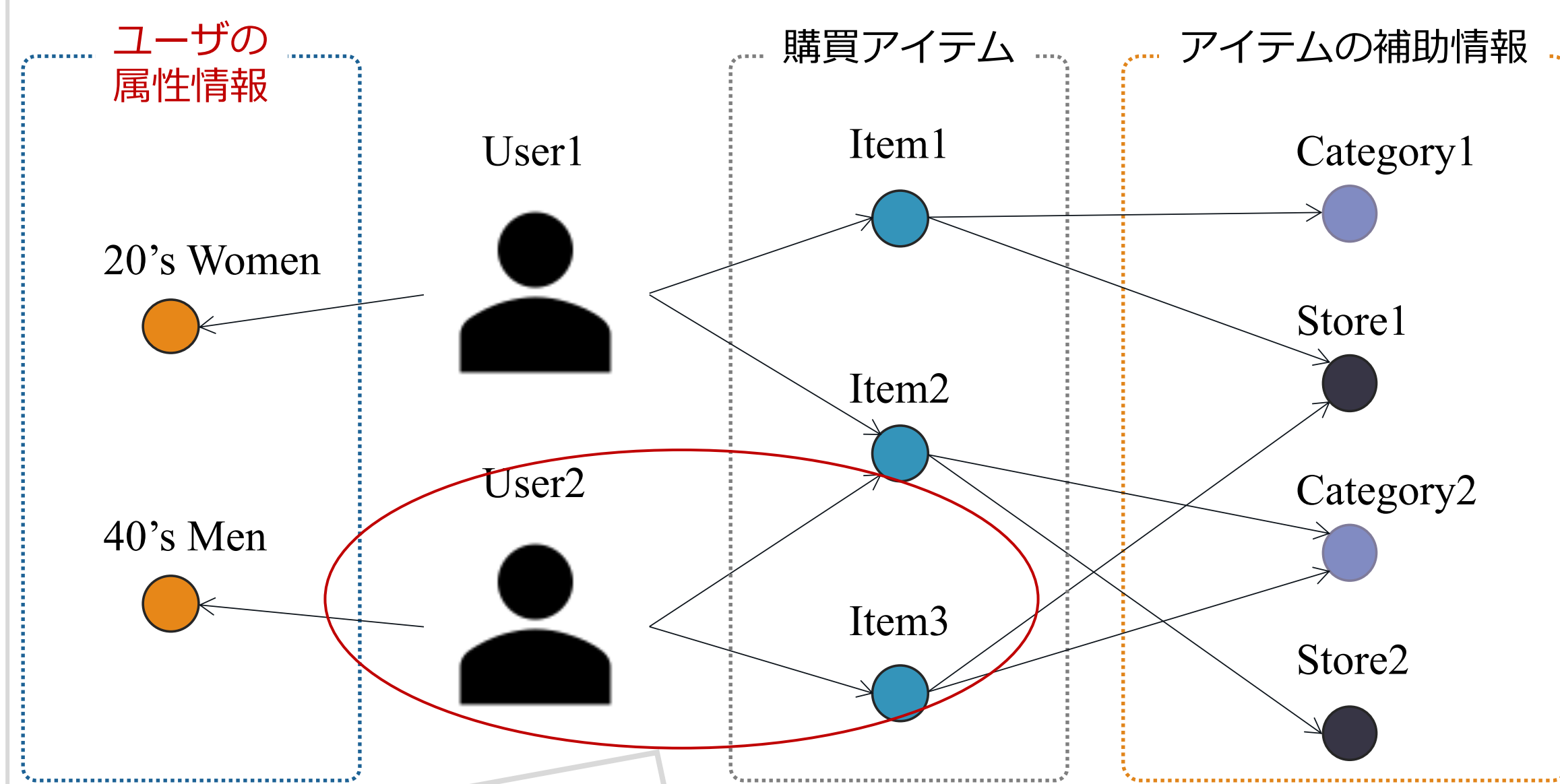
提案モデルの特徴

- グラフ上の各エンティティ、リレーションをベクトル表現で表す
知識グラフを入力とし、各ユーザの各アイテムに対する購買確率を出力する
グラフ上の関係性の重要度をエッジの重みで表現する

提案手法 - グラフの構築

構築するグラフにおいては、以下のエンティティを定義する。

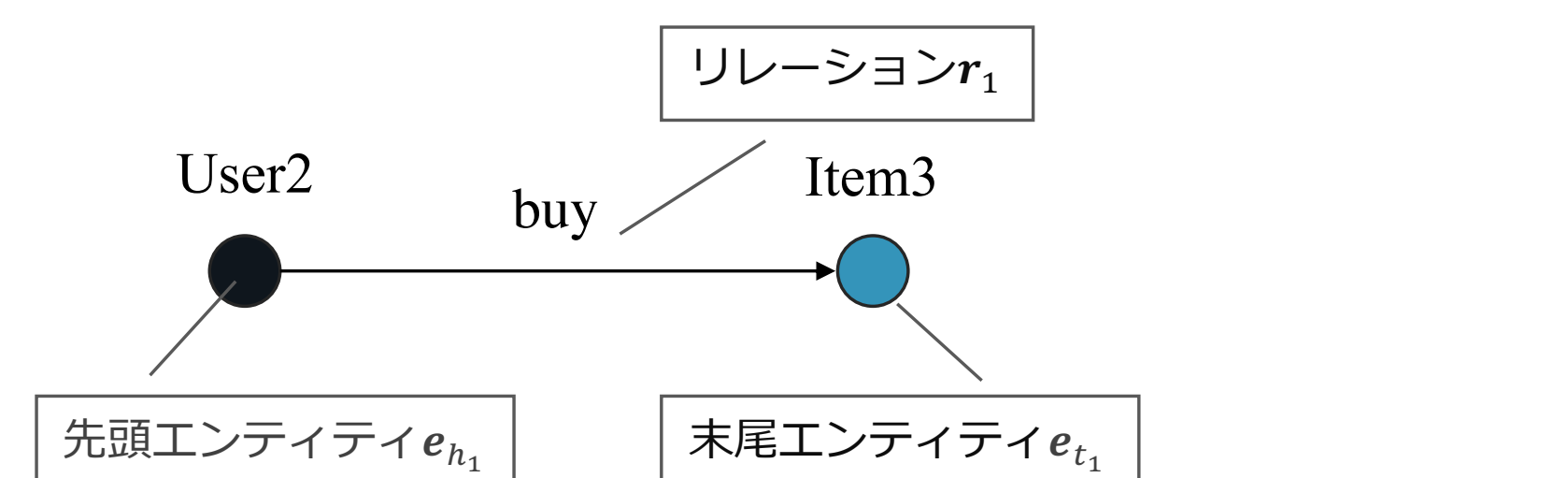
ユーザ集合: U, アイテム集合: I
ユーザ属性集合: P, ストア集合: S, カテゴリ集合: C
エンティティ集合: Q = P U U I U I U H U S U C



知識グラフ (Knowledge Graph)

ノード -> エンティティ, エッジ -> リレーション を表す

先頭エンティティ e\_h, リレーション r, 末尾エンティティ e\_t の3つ組の集合 (トリプル) として定義され、有向グラフで表現される。



提案手法 - モデルの学習

知識グラフの損失とユーザの購買行動に関する損失を同時に最小化することにより最適化を行う

知識グラフの損失: Loss\_KG

グラフで定義されている関係性を正確に表現できているほど小さく

ユーザの購買行動に対する損失: Loss\_CP

ユーザの過去の購買行動を正確に表現できているほど小さく

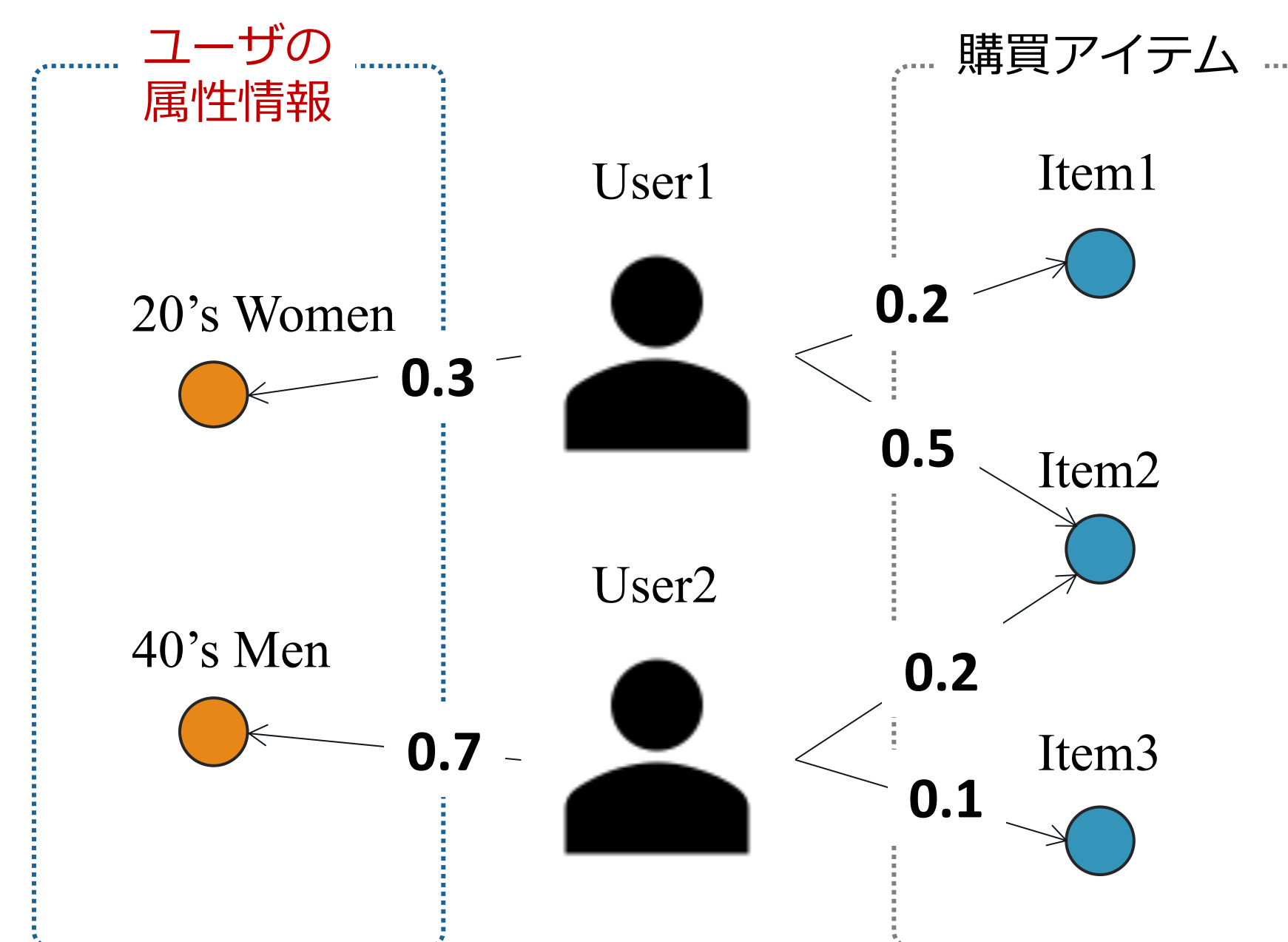
提案手法 - モデルの学習

実際に知識グラフ上で定義されている関係性のうち、ユーザの購買行動を説明する上で重要な関係性にあるエッジの重みが大きくなるような学習方法

モデルの学習により、重み付き有向グラフを得る

エッジの重みを用いた分析を提案する

学習により得られる重み付き有向グラフのイメージ



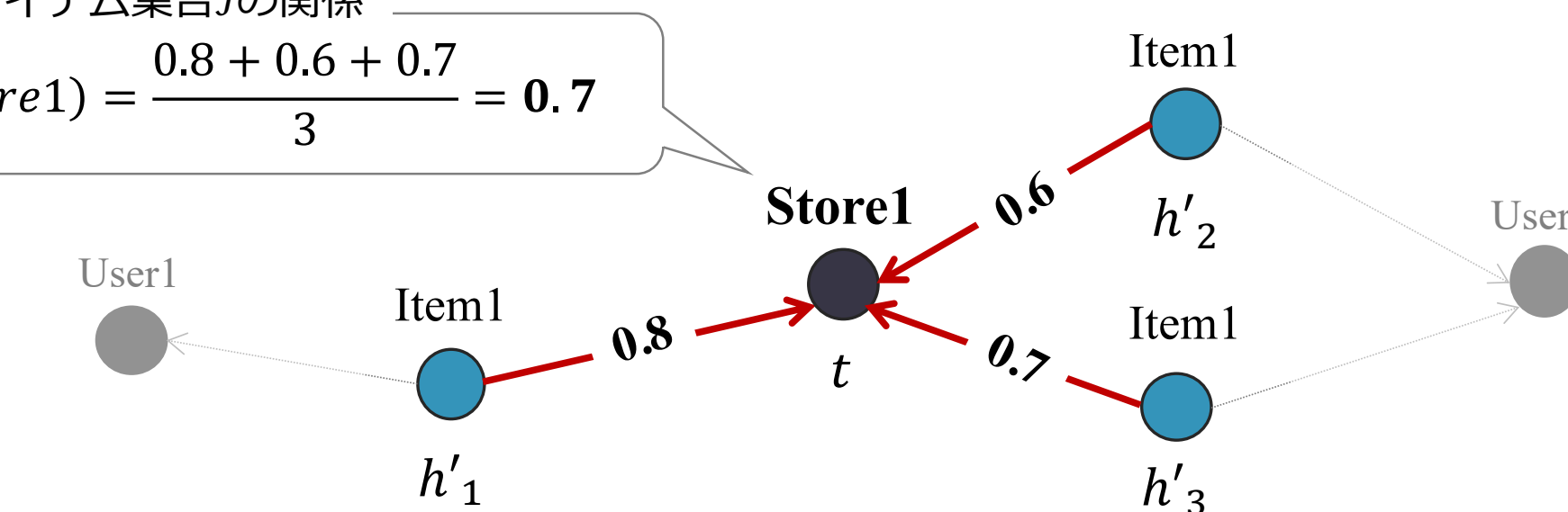
提案手法 - 学習済みモデルを用いた分析

ある末尾エンティティに対して定義されている先頭エンティティのエッジの重みの平均値を算出することで、グラフ上のエンティティ間の関係性の強さを分析する

mu(A, r, t) = sum(h' in A intersect H\_t) pi(h', r, t) / |A intersect H\_t|

A: 任意のエンティティ集合, H\_t subset A: エンティティ t in Q に対する先頭エンティティの集合

Store1とアイテム集合Iの関係
mu(I, r, store1) = (0.8 + 0.6 + 0.7) / 3 = 0.7



実データ分析 - 学習済みモデルを用いた分析

提案手法の有用性を検証するために、楽天市場(2012)の評価履歴データに対して提案手法を適用し分析を行う。

- ユーザの属性: 年代10~70代 (7種) x 性別 (男女) = 14種 (例: 30代男性)
アイテムの補助情報: アイテムの販売ストア, アイテムのカテゴリ
ユーザ数: 27,641, アイテム数: 123,751, カテゴリ数: 13,658, ストア数: 13,123

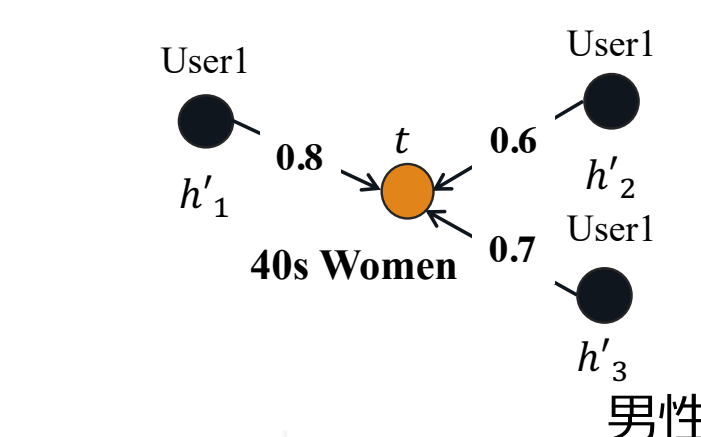
実データ分析 - ユーザと属性の分析

ユーザ集合とユーザ属性 p in P に関して、mu(u, r, p) を算出

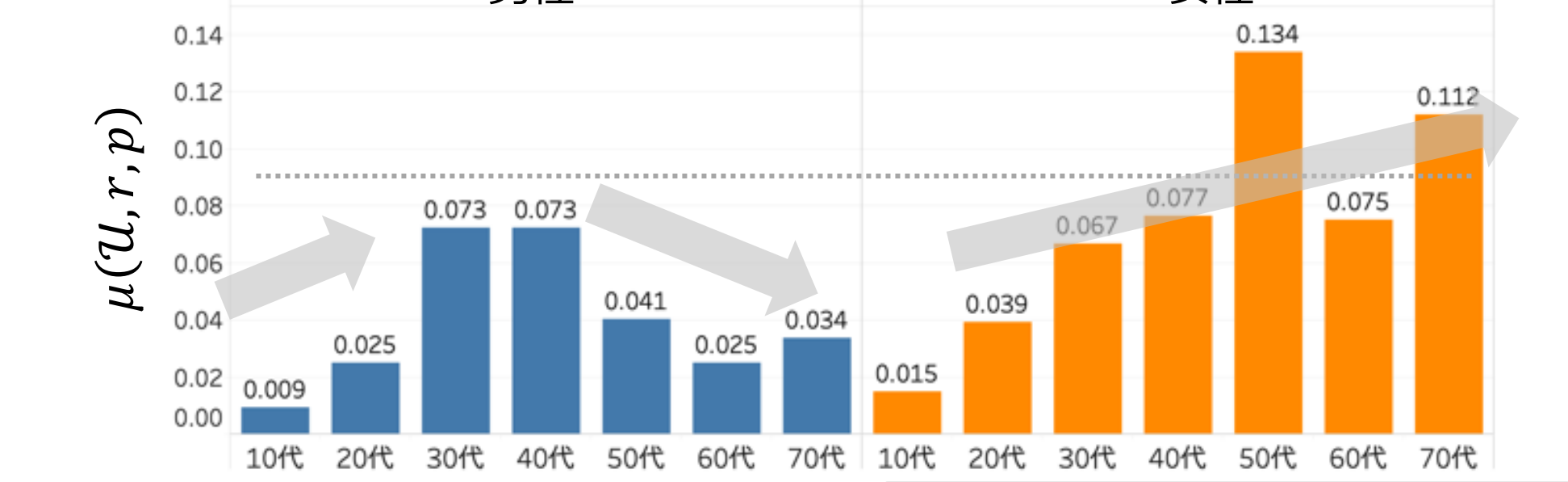
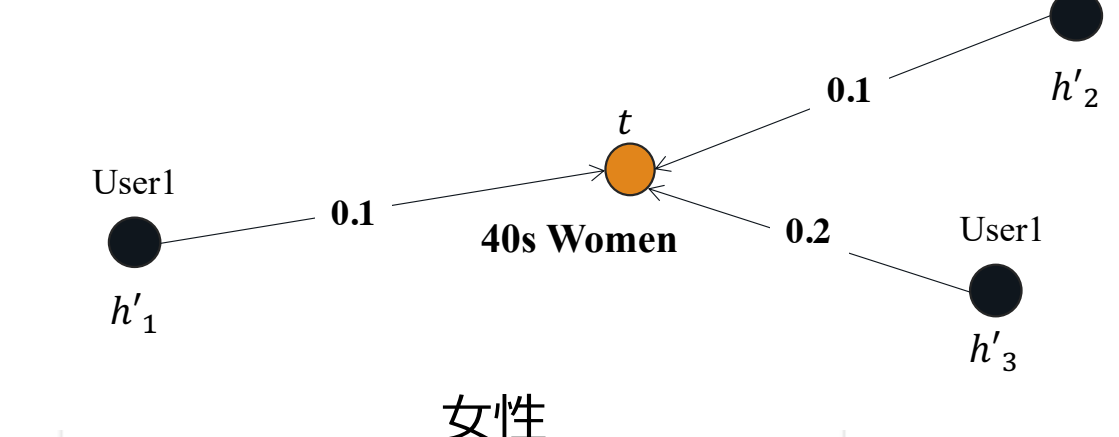
mu(u, r, p) が大きい属性のユーザ同士のベクトルは似た表現を持つ

mu(u, r, p) が小さい属性のユーザ同士のベクトルは異なる表現を持つ

ユーザの嗜好が共通化



ユーザの嗜好が多様化



30代, 40代で最も大きくなり、年代が上昇するにつれて小さくなる

年代が上昇するにつれて大きくなる

50代以降は、趣味などユーザの嗜好の多様化が進む

年齢とともに、嗜好が共通化する

実データ分析 - アイテムとストアの分析

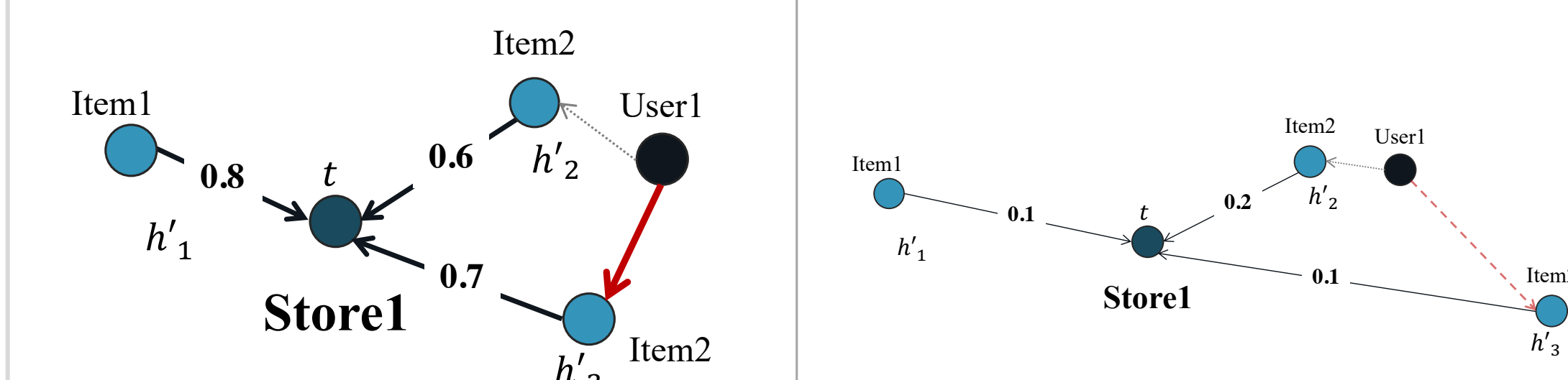
アイテム集合とストア s in S に関して、mu(I, r, s) を算出

mu(I, r, s) が大きい店舗のアイテム同士のベクトルは似た表現を持つ

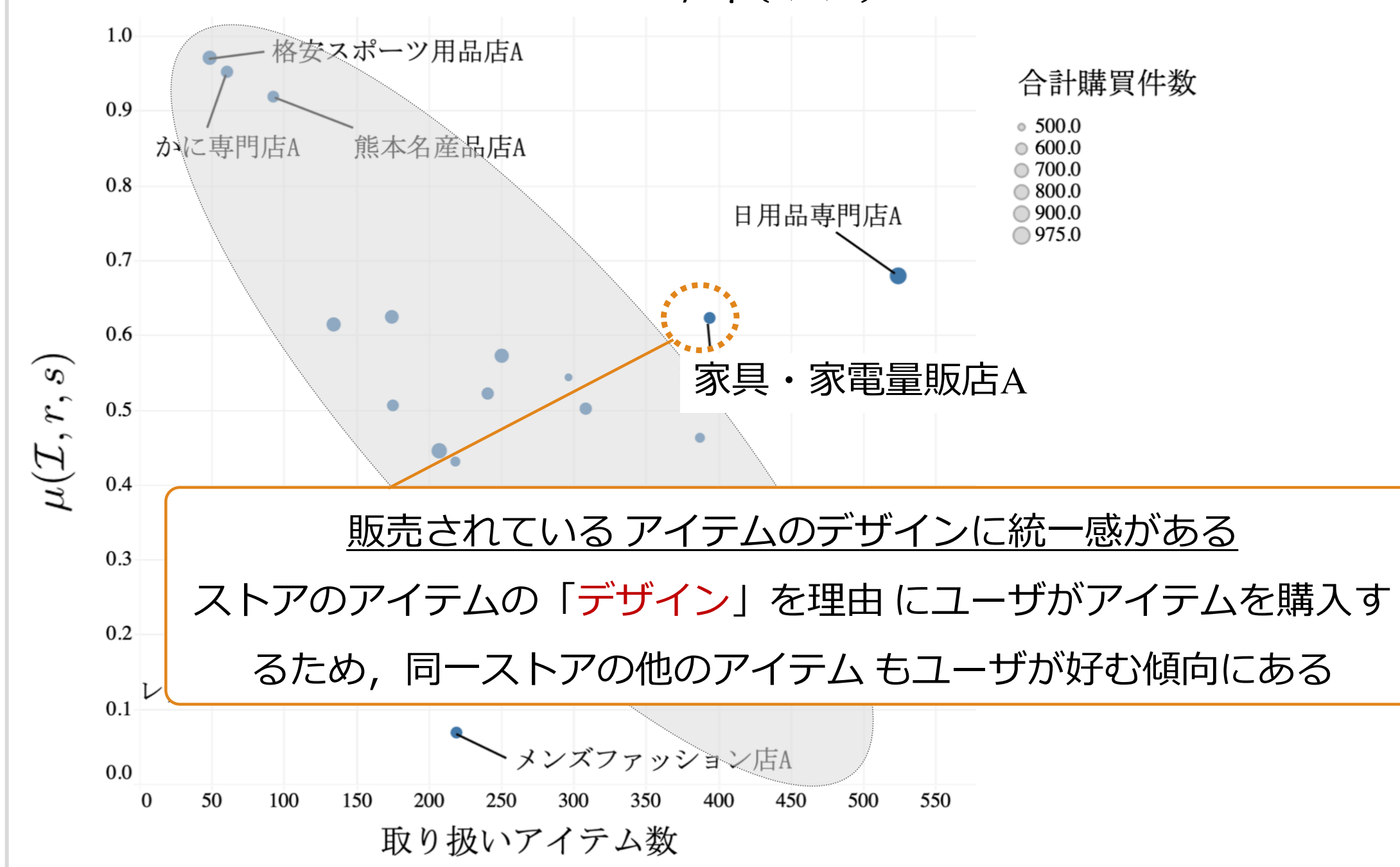
mu(I, r, s) が小さい店舗のアイテム同士のベクトルは異なる表現を持つ

ストアで購入したことのあるユーザが他のアイテムも好む可能性が高い

ストアで購入したことのあるユーザが他のアイテムも好む可能性が低い



取り扱いアイテム数が増加すると mu(I, r, s) は小さくなる傾向にあるが、取り扱いアイテム数が多くても、mu(I, r, s) が大きい店舗も存在する



販売されているアイテムのデザインに統一感があるストアのアイテムの「デザイン」を理由にユーザがアイテムを購入するため、同ストアの他のアイテムもユーザが好む傾向にある

該当店舗の mu(I, r, s) と取り扱いアイテムとの関係