原著論文 Original Paper

次元削減による多項関係予測

A Dimension Reduction Approach to Multinomial Relation Prediction

則 のぞみ
Nozomi Nori
カーネギーメロン大学
Carnegie Mellon University
nozomi.nori@gmail.com

ボレガラ ダヌシカ
Danushka Bollegala
リヴァプール大学
University of Liverpool
danushka.bollegala@liverpool.ac.uk

鹿島 久嗣
Hisashi Kashima
東京大学 / JST さきがけ「知の創生と情報社会」研究領域
The University of Tokyo / JST PRESTO
kashima@mist.i.u-tokyo.ac.jp

keywords: graph, dimension reduction, social media, user modeling, eigenvalue problem

Summary

Many phenomena in the real world can be represented as multinomial relations, which involve multiple and heterogeneous objects. For instance, in social media, users' various actions such as adding annotations to web resources or sharing news with their friends can be represented by multinomial relations which involve multiple and heterogeneous objects such as users, documents, keywords and locations. Predicting multinomial relations would improve many fundamental applications in various domains such as online marketing, social media analyses and drug development. However, the high-dimensional property of such multinomial relations poses one fundamental challenge, that is, predicting multinomial relations with only a limited amount of data. In this paper, we propose a new multinomial relation prediction method, which is robust to data sparsity. We transform each instance of a multinomial relation into a set of binomial relations between the objects and the multinomial relation of the involved objects. We then apply an extension of a low-dimensional embedding technique to these binomial relations, which results in a generalized eigenvalue problem guaranteeing global optimal solutions. We also incorporate attribute information as side information to address the "cold start"problem in multinomial relation prediction. Experiments with various real-world social web service datasets demonstrate that the proposed method is more robust against data sparseness as compared to several existing methods, which can only find sub-optimal solutions.

1. はじめに

1·1 多項関係予測の重要性

データ解析技術の潮流は，個々のデータを対象とした解析から，データ間の関係の解析に移行しつつあると言える．通常のデータ解析では一つのデータについて成り立つ性質を推論するが，データ間の関係の解析ではデータの組（関係データ）について，その間に成立する関係の有無や，関係のもつ性質についての推論を行う．関係データの解析はオンラインマーケティング（顧客と商品の間の購買，評価等の関係），創薬スクリーニング（薬剤と標的の間の関係），ソーシャルネットワーク解析（Facebook, Twitter 等における人間関係）など様々な分野で行われており，データ間の関係に注目することで，個々のデータに注目しているだけでは見えない性質を捉えることができると期待されている．

関係データの中でも最もシンプルなものは，二項関係のデータである．これは二つのオブジェクトの間に特定の関係が成立するかの有無や，その関係の強さについて表現したデータであり，グラフや行列などを用いて容易に表現することができる．ここで，オブジェクトを "関係を構成する，一単位として扱われるデータ"として定義する．近年は，二つのオブジェクト間に限らず，三つ以上の複数のオブジェクト間に生じる "多項関係"に対する解析が盛んである．本論文で定義する多項関係とは "複数種類，複数個のオブジェクト間に生じる一種類の関係" である．例えば，ユーザー A が Web 上のリソース B に対してあるタグ C を付けた行為は，(ユーザー A，リソース B，タグ C）という三種類，三個のオブジェクトの間に生じる一種類（タグ付け）の関係なので，多項関係である*1．多項関係の解析方法は様々あるが，近年盛んに研究されているのがテンソル [Kolda 09] を用いた方法で

*1 また，自然言語処理における述語項構造解析に関連付けて述べると，本論文が対象とする多項関係は，述語 (predicate) が一種類で，項 (argument) が複数種類，複数個にわたる述語項構造を対象としているとも言える．

ある．例えば経時変化するソーシャルネットワークでは，人と人の間の二項関係に対して，その関係がいつ生じたのかという時間情報が加わり三項関係となるが，このような経時変化するネットワークにおける解析は，リンク予測 [Dunlavy 11] や異常検知 [Koutra 12] などのタスクでテンソルを用いて行われている．その他，センサーストリームのデータ解析では（時間，場所，センサーデータの型）の三項関係 [Sun 06a] に対して，ウェブのハイパーリンクの解析では（リンク元ページ，リンク先ページ，アンカーテキスト）の三項関係 [Kolda 05] に対して，インターネットのトラフィック解析では（送信元 IP，送信先 IP，ポートナンバー，時間）の四項関係 [Sun 06b] に対して，それぞれテンソルを用いた解析が行われている．現実のデータに見られるこのような高次の関係（多項関係）を活用することで，二項の関係に注目しているだけでは見えない性質を捉えることができると期待されている．

多種多様な情報源からのデータを統合した解析の重要性が増す中，データ間の高次の関係を捉える多項関係の解析技術は今後益々求められるようになると考えられる．本論文では，とりわけ直接的に意思決定に結びつく予測問題に着目し，多項関係の予測問題を扱う．

1·2 多項関係予測の課題：データ疎性

多項関係予測の重要性は今後増していくと考えられるが，多項関係の予測においては，観測データの疎性に対して頑強な予測を行うことが課題となる．関係データで生じる観測データの疎性の問題とは，関係に含まれるオブジェクト数の増加に伴い，可能な関係の組みわせの数が指数的に増えるため，実際に観測される関係の数が，可能な場合の数に対して相対的に小さくなるために生じるものである．例えば，Cai らの報告によると，広く使用されているソーシャルブックマークサービスのデータセットでは，可能な組み合わせに対して 0.01%程度の関係しか観測されていなかった．また，ほとんどのオブジェクトは少数の関係にしか関与せず，その分布はべき乗分布に従っていた [Cai 11]．後者の疎性は，推薦タスクにおいて重要な問題と見なされているコールドスタート問題 [Schein 02] と呼ばれる状況でしばしば遭遇するものである．これは，ユーザーの過去の行動履歴の情報から推薦を行うシステムにおいて，新規ユーザー（新規アイテム）など，紐付けられる行動履歴の情報が限られたユーザー（アイテム）に対する推薦を行うのが難しくなる問題である．現実の関係データはこのような疎性を示すと考えられるので，このような疎性に対応した予測手法が必要となる．

近年では，多項関係予測を行うにあたってテンソル分解の手法 [Kolda 09, Dunlavy 11, Koutra 12, Sun 06a, Kolda 05, Sun 06b, Symeonidis 08, Rendle 09, Rendle 10] がしばしば利用されているが，多くの手法は非凸最適化問題として定式化されるため，特にデータが疎である場合には局所解による精度悪化が問題となる．近年では凸最適問題として定式化する試みも行われてはいるが，やはり固有値問題を繰り返し解く必要があるため，計算が煩雑になるという問題が存在する．

1·3 提案手法の概要：二項関係への変換，次元削減，属性情報の活用

本論文では，データ疎性に対して頑強であり，かつ，固有値問題を一度解くだけで大域解を得ることができる，多項関係予測のための手法を提案する．我々の提案手法は (1) 多項関係から二項関係の集合への変換，および，それら二項関係集合の低次元空間への埋め込みにより，大域解を保証する最適化問題としての定式化，(2) 未観測のオブジェクトが多数存在する場合でも頑強な予測を行うための，オブジェクト属性の活用という二つのアイデアに基づいている．

まず，K 個のオブジェクトを巻き込む関係を，各オブジェクトと，そのオブジェクトを巻き込む関係インスタンスの間の K 個の二項関係に変換する．ここで，関係インスタンスは“オブジェクト間に生じた関係それ自体を一単位としたデータ”と定義する．図 1 に，(花子 , www.ai-gakkai.or.jp, 太郎) の間の多項関係を，関係インスタンスと，“花子”，“ ai-gakkai.or.jp”，“太郎”の三つのオブジェクトの間の二項関係の集合に変換した例を示す．このような変換はハイパーグラフの接続行列 [Voloshin 09] に相当する．続いて，上記の変換で得られた二項関係の集合に対して，非線形次元削減手法 [Belkin 03] を適用することで，各オブジェクトとそのオブジェクトが参加した関係インスタンスが潜在空間で近傍に位置するような潜在空間への写像を学習し，異種のオブジェクトを共通の潜在空間に埋め込む．結果として得られる最適化問題は，大域解を保証する一般化固有値問題として定式化されるため，データ疎性に対して頑強な予測が可能となる．加えて，多項関係予測におけるコールドスタート問題に対処するために，異種オブジェクトの様々な属性を活用する．例えば，オブジェクトとしてユーザを考えると，ユーザは，年齢，性別，所属などで表現することができるだろう．このようなオブジェクトの属性は，ほとんどのオブジェクトが僅かな関係にしか関与しないような状況では予測に有用であると期待できる．

我々は，提案手法の頑強性を評価するために，現実のデータセットとして，ソーシャルメディア上でのユーザーの行動データを用いて実験を行った．結果，提案手法は，(1) 訓練時に少数の関係データしか得られない状況や，(2) 予測時に大量の未知のオブジェクトが存在するような状況といった，データの疎性が顕著な状況下で，標準的なテンソル分解手法を予測精度で上回り，データが疎な状況でも安定した予測精度を実現した．

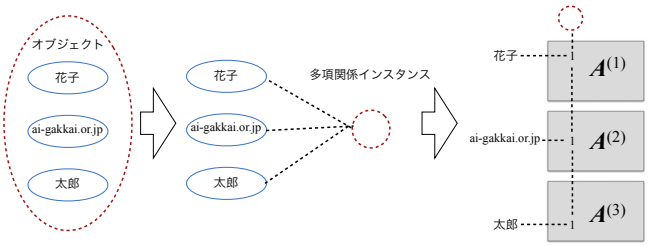


図 1 三種類 ($K=3$) のオブジェクトを巻き込む一つの多項関係を三つの二項関係に変換し，更にそれらが，接続行列における要素として表現されることを示す例．

2. 提 案 手 法

2·1 問 題 設 定

人や Web ページなど，異なる種類のオブジェクトの間で，ある特定の種類の関係がどれくらい生じやすそうかを，いくつかの観測された関係データを元に推定する問題を考える．例えば，ある人が，別の誰かから薦められた Web ページを気に入るかを予測する状況を考えてみよう．ここでの目標は，“$\texttt{person}_1$ が $\texttt{person}_2$ によって薦められた URL で指定される Web ページを気に入る”という関係が，($\texttt{person}_1$, URL, $\texttt{person}_2$) の各組み合わせについてどれくらい生じやすいかを (花子 , www.ai-gakkai.or.jp, 太郎) などの既知の事実を元に予測することである．

K 種類のオブジェクト集合，$S^{(1)}, S^{(2)}, \ldots, S^{(K)}$ があり，それぞれの集合が $N^{(k)}$ $(1 \leq k \leq K)$ 個のオブジェクトを持つとする．先の例では，$S^{(1)}$ と $S^{(3)}$ をそれぞれユーザの集合に，$S^{(2)}$ を URL の集合として取ることができる．i 番目のオブジェクト $s^{(k,i)} \in S^{(k)}$ を，$s^{(k,i)}$ のように記載する．例えば，$s^{(1,1)}$ によって花子 を表す．また，M 個の観測された関係インスタンスから成る集合，$O \subset S^{(1)} \times S^{(2)} \times \cdots \times S^{(K)}$ が得られているとする．各関係インスタンスは，ある特定の関係（例えば先の例では，気に入るという関係）がオブジェクトの特定の組み合わせに対して成立していることを示している．例えば，$o^{(1)} \in O$ は (花子 , www.ai-gakkai.or.jp, 太郎) などである．

さて，我々の目標はオブジェクトの各組み合わせの中で，観測された関係インスタンスの集合 O に含まれていないものについて，特定の関係がどれくらい生じやすいかを予測することである．

多くの現実的な状況では，各オブジェクトは自身に関する何かしらの情報を有する．例えば，人であれば年齢や性別などのデモグラフィックな情報を持つと期待できる．したがって，$s^{(k,i)}$ に対して，$D^{(k)}$ 次元の属性ベクトル $\boldsymbol{x}^{(k,i)}$ を関連付け，各 $k = 1, 2, \ldots, K$ について，まとめて計画行列

$$\boldsymbol{\Phi}^{(k)} \equiv (\boldsymbol{x}^{(k,1)}, \boldsymbol{x}^{(k,2)}, \ldots, \boldsymbol{x}^{(k,N^{(k)})})^\top$$

とする．

本論文が提案する多項関係予測の問題は，以下のような入出力を持つ問題として要約できる．

問題：多項関係予測

- **入力:**
  - $S^{(1)}, S^{(2)}, \ldots, S^{(K)}$: K 種類のオブジェクト集合
  - O $(\subset S^{(1)} \times S^{(2)} \times \cdots \times \ldots S^{(K)})$: 観測された M 個の関係インスタンスから成る集合
  - $\boldsymbol{\Phi}^{(1)}, \boldsymbol{\Phi}^{(2)}, \ldots, \boldsymbol{\Phi}^{(K)}$: オブジェクトの属性を表現する K 個の計画行列
- **出力:** O に含まれない，すなわち，$(S^{(1)} \times S^{(2)} \times \cdots \times \ldots S^{(K)}) \setminus O$ に含まれるオブジェクトの組み合わせに対して，関係の生じやすさを表すスコア．

2·2 次元削減を用いた多項関係予測

ここでは，次元削減手法を用いて，一般化固有値問題を一度解くだけで大域解が求まる，新しい多項関係予測手法を提案する．まず，オブジェクトの属性を考慮しない場合から考える．

大域解を保証するための最初のキーアイデアは，多項関係を二項関係に変換することである．K 種類のオブジェクト集合それぞれに対して，一つの二値行列を構築し，全部で K 個の行列を得る．これらの行列は“オブジェクトと関係インスタンスの間の関係”を表現するものである．行列の各要素は，ある特定のオブジェクトが，ある特定の関係インスタンスに参加しているかを示している．

$\boldsymbol{A}^{(k)}$ を，$S^{(k)}$ に属するオブジェクトの，O に属する関係インスタンスに対する参加情報をまとめた $N^{(k)} \times M$ の二値行列とする．$\boldsymbol{A}^{(k)}$ の各要素 $[\boldsymbol{A}^{(k)}]_{n,m}$ は，以下のように定義される．

$$[\boldsymbol{A}^{(k)}]_{n,m} \equiv \begin{cases} 1 & (s^{(k,n)} \in S^{(k)} \text{ が } o^{(m)} \in O \text{に参加する場合}) \\ 0 & (\text{そうでない場合}). \end{cases}$$

図 1 にこの変換の例を示している．

我々の二つ目のキーアイデアは，K 個の行列によって表現された二項関係を低次元へ埋め込むことである．次元削減を用いた二部グラフ予測の手法 [Yamanishi 09] と似たアイデアを用いて，各オブジェクトとそのオブジェクトが参加した関係インスタンスが潜在空間で近傍に位置するような潜在空間への写像を学習し，異種のオブジェクトと関係インスタンスを共通の潜在次元に埋め込む．

まず最初に一次元の潜在空間への埋め込みを考えてみよう．サイズ $N^{(1)}$ のオブジェクト集合 $S^{(1)}$ は，長さ $N^{(1)}$ のベクトル $\boldsymbol{f}^{(1)}$ として埋め込まれる．同様にして，各オブジェクト集合 $S^{(2)}, S^{(2)}, \ldots, S^{(K)}$ は，それぞれ $\boldsymbol{f}^{(2)}, \boldsymbol{f}^{(3)}, \ldots, \boldsymbol{f}^{(K)}$ として埋め込まれる．サイズ M の観測された関係インスタンスの集合 O もまた，同じ一次元の潜在空間に，長さ M のベクトル $\bar{\boldsymbol{f}}$ として埋め込まれる．

もし，あるオブジェクト $s^{(k,n)} \in S^{(k)}$ が，ある関係インスタンス $o^{(m)} \in O$ に参加するならば，両者の埋め込み先，$[\boldsymbol{f}^{(k)}]_n$ と $[\bar{\boldsymbol{f}}]_m$ を近くすること，すなわち，ユークリッド距離 $([\boldsymbol{f}^{(k)}]_n - [\bar{\boldsymbol{f}}]_m)^2$ を小さくすることを試みよう．このとき，最終的に最小化すべき目的関数は以下のように定義される．

$$
\begin{aligned}
&J(\{\boldsymbol{f}^{(k)}\}_{k=1}^K, \bar{\boldsymbol{f}}) \qquad (1)\\
&= \sum_k \sum_i \sum_j [\boldsymbol{A}^{(k)}]_{i,j} \left([\boldsymbol{f}^{(k)}]_i - [\bar{\boldsymbol{f}}]_j \right)^2 \\
&= \sum_k \left({\boldsymbol{f}^{(k)}}^\top \boldsymbol{D}^{(k)} \boldsymbol{f}^{(k)} + \bar{\boldsymbol{f}}^\top \bar{\boldsymbol{f}} - 2 {\boldsymbol{f}^{(k)}}^\top \boldsymbol{A}^{(k)} \bar{\boldsymbol{f}} \right),
\end{aligned}
$$

ここで，$\boldsymbol{D}^{(k)}$ は，その (i,i) 番目の要素が $[\boldsymbol{D}^{(k)}]_{i,i} \equiv \sum_j [\boldsymbol{A}^{(k)}]_{i,j}$，すなわち，オブジェクト $s^{(k,i)}$ が参加した関係の数として定義される対角行列である．ここで，各 j について $\sum_i [\boldsymbol{A}^{(k)}]_{i,j} = 1$ が成立することを用いた．

この目的関数は，$\boldsymbol{f}^{(k)} \equiv \boldsymbol{0}$ かつ $\bar{\boldsymbol{f}} \equiv \boldsymbol{0}$ とすることで容易に最小化できてしまうので，このような望ましくない解を避けるために，以下のスケーリング制約を加える．

$$
\sum_{k=1}^K {\boldsymbol{f}^{(k)}}^\top \boldsymbol{D}^{(k)} \boldsymbol{f}^{(k)} = 1. \qquad (2)
$$

目的関数 (1) の $\bar{\boldsymbol{f}}$ に関する最小値は $\frac{\partial J}{\partial \bar{\boldsymbol{f}}} = 0$ とおいて以下のように得られる．

$$
\bar{\boldsymbol{f}} = \frac{1}{K} \sum_{k=1}^K {\boldsymbol{A}^{(k)}}^\top \boldsymbol{f}^{(k)}. \qquad (3)
$$

式 (3) を式 (1) の正負を逆転させたものに代入することで，以下の最大化問題を得る．

$$
\begin{aligned}
&-J(\{\boldsymbol{f}^{(k)}\}_{k=1}^K) \qquad (4)\\
&= \frac{1}{K} \sum_{k,\ell=1}^K {\boldsymbol{f}^{(k)}}^\top \boldsymbol{A}^{(k)} {\boldsymbol{A}^{(\ell)}}^\top \boldsymbol{f}^{(\ell)} - \sum_{k=1}^K {\boldsymbol{f}^{(k)}}^\top \boldsymbol{D}^{(k)} \boldsymbol{f}^{(k)}.
\end{aligned}
$$

ゆえに，

$$
\begin{aligned}
&L(\{\boldsymbol{f}^{(k)}\}_{k=1}^K, \lambda) = \qquad (5)\\
&\quad -J(\{\boldsymbol{f}^{(k)}\}_{k=1}^K) - \lambda \left(\sum_{k=1}^K {\boldsymbol{f}^{(k)}}^\top \boldsymbol{D}^{(k)} \boldsymbol{f}^{(k)} - 1 \right),
\end{aligned}
$$

として定義されるラグランジュ関数を，$\frac{\partial L}{\partial \boldsymbol{f}^{(k)}} = 0, \frac{\partial L}{\partial \lambda} = 0$ とおいて最大化することで以下を得る．

$$
\sum_\ell \boldsymbol{A}^{(k)} {\boldsymbol{A}^{(\ell)}}^\top \boldsymbol{f}^{(\ell)} = K(\lambda+1) \boldsymbol{D}^{(k)} \boldsymbol{f}^{(k)}. \qquad (6)
$$

さらに，$\tilde{\lambda} \equiv K(\lambda+1)$ とすることで，以下の一般化固有値問題を得ることができる．

$$
\boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{f} = \tilde{\lambda} \boldsymbol{D} \boldsymbol{f}, \qquad (7)
$$

ここで，$\boldsymbol{A}$，$\boldsymbol{D}$ と $\boldsymbol{f}$ は以下のように定義される．

$$
\begin{aligned}
\boldsymbol{A} &\equiv \begin{bmatrix} \boldsymbol{A}^{(1)} \\ \boldsymbol{A}^{(2)} \\ \vdots \\ \boldsymbol{A}^{(K)} \end{bmatrix} \\
\boldsymbol{D} &\equiv \begin{bmatrix} \boldsymbol{D}^{(1)} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{D}^{(K)} \end{bmatrix} \\
\boldsymbol{f} &\equiv ({\boldsymbol{f}^{(1)}}^\top, {\boldsymbol{f}^{(2)}}^\top, \ldots, {\boldsymbol{f}^{(K)}}^\top)^\top.
\end{aligned}
$$

最大固有値に対応する固有ベクトル $\boldsymbol{f}$ が，オブジェクトの，一次元空間での最適な埋め込み先である．R 次元の空間における埋め込み先 $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_R$ を得るためには，固有値の大きい順に上位 R 個の固有ベクトルを取得すれば良い．

最後に，オブジェクトのある組み合わせ $o \equiv (s^{(1,i_1)}, s^{(2,i_2)}, \ldots, s^{(K,i_K)})$ に対して，それぞれのオブジェクトの埋め込み先 $[\boldsymbol{f}^{(k)}]_{i_k}$ と，その組み合わせの最適な埋め込み先 $\bar{\boldsymbol{f}}$ の間にも式 (3) は成立するので，o の最適な埋め込み先の r 番目の次元は

$$
\frac{1}{K} \sum_{k=1}^K [\boldsymbol{f}_r^{(k)}]_{i_k} \qquad (8)
$$

で与えられる．また，式 (1) で定義したように，オブジェクトと関係インスタンスの間のユークリッド距離が小さいほど，そのオブジェクトと関係インスタンスの間に関係が生じやすいので，関係 o は，

$$
\mathit{diff}(o) \equiv \sum_{r=1}^R \sum_{k=1}^K \left([\boldsymbol{f}_r^{(k)}]_{i_k} - \frac{1}{K} \sum_{k'=1}^K [\boldsymbol{f}_r^{(k')}]_{i_{k'}} \right)^2. \qquad (9)
$$

の値が小さいほど生じやすいと考えられる．ゆえに，O に含まれない組み合わせ o に対して，関係の生じやすさを表すスコア $\mathit{score}(o)$ は $\mathit{score}(o) \equiv -\mathit{diff}(o)$ などとして与えられる．

非凸な目的関数を持つ当初の最適化問題 (5) が一般化固有値問題に帰着されたことは特筆すべきである．一般化固有値問題は厳密に解くことができるため，当初の最適化問題に対して大域解を得ることが可能になる．この性質は，局所解のみ保証する既存の多くのテンソル分解手法と対照的である．更に，既存の多くの手法 [Kolda 09] では，複数の固有値問題を解かなければならない一方，提案手法では一般化固有値問題を一度解くだけで大域解が求まる．

2·3 オブジェクトの属性の活用

続いて，オブジェクトの属性情報 $\{\boldsymbol{\Phi}^{(k)}\}_{k=1}^{K}$ を統合することを考える．オブジェクトの属性を活用することは，僅かな関係にしか参加しないようなオブジェクトが存在する場合や，予測を，訓練段階では未観測であったオブジェクトに対して行う必要がある場合などにおいて特に重要となる．

線形写像

$$\boldsymbol{f}^{(k)} \equiv \boldsymbol{\Phi}^{(k)}\boldsymbol{w}^{(k)},$$

を考える．ここで，$\boldsymbol{w}^{(k)}$ は $D^{(k)}$ 次元の属性ベクトルを一次元の潜在空間に写像する $D^{(k)}$ 次元のパラメータである．

属性情報を活用しない場合と同様にラグランジュ関数を定義し，それを最適化することで，

$$\sum_{\ell} \boldsymbol{\Phi}^{(k)\top}\boldsymbol{A}^{(k)}\boldsymbol{A}^{(\ell)\top}\boldsymbol{\Phi}^{(\ell)}\boldsymbol{w}^{(\ell)} = K\,(\lambda+1)\,\boldsymbol{\Phi}^{(k)\top}\boldsymbol{D}^{(k)}\boldsymbol{\Phi}^{(k)}\boldsymbol{w}^{(k)}, \quad (10)$$

を得ることができる．これは以下の一般化固有値問題として表現できる．

$$\boldsymbol{\Phi}^{\top}\boldsymbol{A}\boldsymbol{A}^{\top}\boldsymbol{\Phi}\boldsymbol{w} = \tilde{\lambda}\boldsymbol{\Phi}^{\top}\boldsymbol{D}\boldsymbol{\Phi}\boldsymbol{w}, \quad (11)$$

ここで，

$$\boldsymbol{\Phi} \equiv \begin{bmatrix} \boldsymbol{\Phi}^{(1)} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{\Phi}^{(K)} \end{bmatrix}$$

$$\boldsymbol{w} \equiv (\boldsymbol{w}^{(1)\top}, \boldsymbol{w}^{(2)\top}, \ldots, \boldsymbol{w}^{(K)\top})^{\top}.$$

である．属性ベクトルの次元が高い場合，次元の呪いと呼ばれる効果により予測性能がしばしば悪化する．これを避けるため，正の値である正則化パラメータ $\sigma > 0$ により，正則化項を追加するのが一般的である．この場合，我々の一般化固有値問題 (11) は以下のように修正される．

$$\boldsymbol{\Phi}^{\top}\boldsymbol{A}\boldsymbol{A}^{\top}\boldsymbol{\Phi}\boldsymbol{w} = \lambda\left(\boldsymbol{\Phi}^{\top}\boldsymbol{D}\boldsymbol{\Phi} + \sigma\boldsymbol{I}\right)\boldsymbol{w}. \quad (12)$$

2·4 計算量について

提案手法の処理は，一般化固有値問題を解きオブジェクトの潜在次元表現を得る次元削減部分と，その潜在次元を用いた予測部分の二つに分けられる．前者については，属性を考慮した式 (12) で $\boldsymbol{\Phi}$ を単位行列として $\sigma = 0$ とおけば式 (7) になるので，式 (12) について述べると，この計算量は $\boldsymbol{B} = \boldsymbol{\Phi}^{\top}\boldsymbol{A}\boldsymbol{A}^{\top}\boldsymbol{\Phi}$, $\boldsymbol{C} = \boldsymbol{\Phi}^{\top}\boldsymbol{D}\boldsymbol{\Phi} + \sigma\boldsymbol{I}$ とおいたときの一般化固有値問題 $\boldsymbol{B}\boldsymbol{x} = \lambda\boldsymbol{C}\boldsymbol{x}$ の計算量に相当する．今，$\boldsymbol{C}$ は正定値対称行列であり，原理的に $\boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{x} = \lambda\boldsymbol{x}$ と固有値問題に帰着できるが，本論文では $\boldsymbol{C} = \boldsymbol{L}\boldsymbol{L}^{T}$（$\boldsymbol{L}$ は下三角行列）と $\boldsymbol{C}$ をコレスキー分解し，$\boldsymbol{F}\boldsymbol{y} = \lambda\boldsymbol{y}, (\boldsymbol{F} = \boldsymbol{L}^{-1}\boldsymbol{B}\boldsymbol{L}^{-T}, \boldsymbol{y} = \boldsymbol{L}^{T}\boldsymbol{x})$ という固有値問題に帰着させる．$\boldsymbol{B}, \boldsymbol{C}$ を $n * n$ の行列とすると，$\boldsymbol{C}$ のコレスキー分解と $\boldsymbol{L}, \boldsymbol{L}^{T}$ の逆行列化の計算量は，並列化なしの場合はいずれも $O(n^3)$ だが，並列化によってそれぞれ $O(\log^3 n), O(\log^2 n)$ となる [Csanky 76, Pan 87]．帰着された固有値問題 $\boldsymbol{F}\boldsymbol{y} = \lambda\boldsymbol{y}$ においては，$\boldsymbol{F}$ は対称行列なので，Implicitly Restarted Lanczos Method (IRLM)[Lehoucq 96, Sorensen 92] を用いて固有値問題を解くことができ，この計算量は $O(mRh + nR^2h)$ (ここで，n と m はそれぞれ 行列 $\boldsymbol{F}$ に対応するグラフのノード数とエッジ数，h は IRLM が収束するまでのイテレーション回数，R は潜在次元数) となる [White 05]．これは，n, m についてそれぞれ線形である．

ある関係インスタンスが与えられたときにスコアを付与する計算量は，式 (9) の計算量として $O(K)$(K は項数) となる（次元数 R に関する和を取る部分はベクトルとして表記可能なため）．オブジェクトの全組み合わせに対して式 (9) を計算するには，現状では全組み合わせを列挙するため $O(\sqcap_{k=1}^{K} N^{(k)})$ となる．ここで，$N^{(k)}$ は，k 種類目のオブジェクトのオブジェクト数である．実用的には，スコアの高いもの上位 l 個についてその組み合わせとスコアを得られれば十分な場合も想定される．そのようなアプリケーションに応じた対応も含め，計算量の削減が課題となる．

3. 実　　験

ここでは，ソーシャルメディア上でのユーザーの行動予測を応用として，提案手法がデータ疎性に対する高い頑健性を持つことを，現実の三種類のデータセットを用いた実験によって示す．

3·1 応用事例：ソーシャルメディア上でのユーザの行動予測

Facebook や Twitter，Tumblr などに代表されるソーシャル Web サービスは，検索サービスと並び，World Wide Web の主要な利用の場として普及した．ソーシャル Web サービスにおける重要なコンセプトの一つはユーザの行動である．例えば，ユーザは，ソーシャルブックマークサービスを使い，Web ページ，写真，論文などの様々なリソースに対してキーワードを付与することができ，更にそれを他のユーザと共有することができる．ユーザはまた，Twitter における “retweet”や Tumblr における “reblog”などの機能を通じて，他のユーザーが発信した情報を，自身のソーシャルネットワークを介して他のユーザに再発信することができる．近年，これらの行為データを Web における推薦や，個人化された検索などの様々なタスクに活用する試みが盛んに行われている．例えば，ユーザのタグ付けデータは，検索 [Bao 07, Heymann 08]，人間関係の推論 [Schifanella 10]，オントロジーの発見 [Mika

05] などのタスクに有用であることが示されている．更に，ユーザの行動自体の予測が可能になるとその応用可能性は一段と広がる．一例として，ソーシャルブックマークサービス上でユーザが各リソースに対して付与できるタグを推薦する，タグ推薦のタスクが挙げられる．タグ推薦の機能によって，より多くのユーザがより多くのリソースに対してタグを付与するようになることで，リソースに付随する情報を増やすことができ，情報抽出の質を向上させることができる [Guan 09] と期待できる．本実験では，deli.cio.us と呼ばれるソーシャルブックマーク上でのタグ付け行為と，マイクロブログサービス Twitter 上でのユーザーのいくつかの行為を対象として，その行動予測を行う．データの詳細は以下で述べる．

3·2 データセット

表 1 に，使用した三つのデータセットについて詳細を記載した．最初の二つのデータセット[*2]は，マイクロブログサービス Twitter から取得したものであり，ここでは “retweet”と “favorite”という二つの行為を対象とした．各行為は，行為主体ユーザ (subjective user)，発信元ユーザ (mentioned user)，URL という，三種類のオブジェクトから構成されるタプルとして表現できる．各タプルにより表現されているのは，発信元ユーザの tweet によって投稿された URL に対して，行為主体ユーザが特定の行為（retweet/favorite）を行ったということである．三つ目のデータセット[*3]は del.icio.us というソーシャルタギングサービスから取得したものであり，各行為は，ユーザ (user)，ユーザによって付与されたタグ (tag)，URL から成るタプルで表現される．提案手法では，関係データの他に表 1 に示されるような各オブジェクトの属性も活用した．詳細は表 1 にまとめられている．特徴ベクトルを構築する際には TF-IDF を用い，属性値が $[0.0, 1.0]$ の範囲に収まるようにした．deli.cio.us における “friend 関係にあるユーザ”の属性以外は，各オブジェクトについて TF-IDF 値が上位 5 つの属性値を持つ属性を活用した．

3·3 実験設定：関係レベルで疎な状況とオブジェクトレベルで疎な状況

実験条件として，導入で述べた，多項関係予測において問題となる二種類の疎な状況，すなわち，関係レベルで疎である状況とオブジェクトレベルで疎である状況を設定した．関係レベルで疎な設定は，いくつかの関係インスタンスがランダムに欠けている状況を仮定する．オブジェクトレベルで疎な設定は，特定のオブジェクトを含む全ての関係インスタンスが欠けている状況を仮定する．後者の疎な状況は，推薦システムにおけるコールドスタート問題として知られている．例えば，新しいユーザが初めてオンラインショッピングのサイトに訪れた時，大抵の場合，そのユーザに関して得られる情報はほとんど，もしくは全くない．このため，そのユーザに対して何らかの行動予測を行うことは難しい．

関係レベルで疎な設定では，全データセットから，観測された関係インスタンスの一定割合をランダムにサンプリングし訓練データとして用い，残りのデータを評価データとした．オブジェクトレベルで疎な設定では，観測されたオブジェクトの一定割合をランダムにサンプリングし，そのオブジェクトを含まなかった関係インスタンスを訓練データとして用いた．残りのデータは評価データとして用いた．それぞれの実験設定で，サンプリング比率を変え，各サンプリング比率について，サンプリング，予測，評価の一連のプロセス（以降，この一連のプロセスをスロットと呼ぶ）を 10 回繰り返した．予測性能の評価指標としては AUC を用いた．AUC は，評価データの中で，未観測の関係インスタンスのリンク強度よりも，観測済みの関係インスタンスのリンク強度が高い値となる確率として計算できる．

3·4 比較手法

比較手法としては，CP 分解 (PARAFAC) と，Tucker 分解と呼ばれる二つの標準的なテンソル分解の手法を採用した．CP 分解/Tucker 分解は，元のテンソルと，近似したテンソルの間に定義した距離を最小化する最適化問題として定式化される．CP 分解は，テンソルをランク 1 のテンソルの和として近似するものであり，Tucker 分解は，テンソルをコアテンソルといくつかの因子行列を用いて近似する．CP 分解では，三階テンソル $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$ が与えられた時，このテンソルは $\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \mathbf{a_r} \circ \mathbf{b_r} \circ \mathbf{c_r}$ によって近似される．ここで，$r = 1, \ldots, R$ について $\mathbf{a_r} \in \mathbb{R}^I, \mathbf{b_r} \in \mathbb{R}^J, \mathbf{c_r} \in \mathbb{R}^K$ であり，$\circ$ は外積を表す．$\boldsymbol{\mathcal{X}}$ の各要素 $[\boldsymbol{\mathcal{X}}]_{i,j,k}$ は，$i = 1, \ldots, I, j = 1, \ldots, J, k = 1, \ldots, K$ について $[\boldsymbol{\mathcal{X}}]_{i,j,k} \approx \sum_{r=1}^{R} [\mathbf{a_r}]_i [\mathbf{b_r}]_j [\mathbf{c_r}]_k$ として近似される．Tucker 分解では，$\boldsymbol{\mathcal{X}}$ の各要素 $[\boldsymbol{\mathcal{X}}]_{i,j,k}$ は，

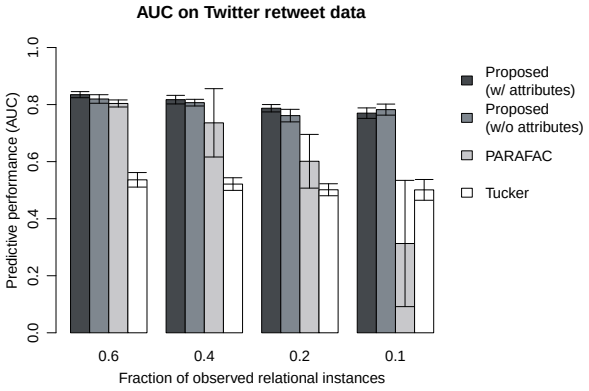
$$[\boldsymbol{\mathcal{X}}]_{i,j,k} \approx \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} [\boldsymbol{\mathcal{G}}]_{p,q,r} [\mathbf{A}]_{i,p} [\mathbf{B}]_{j,q} [\mathbf{C}]_{k,r}$$

によって近似される．$\boldsymbol{\mathcal{G}} \in \mathbb{R}^{P \times Q \times R}$ はコアテンソル，$\mathbf{A} \in \mathbb{R}^{I \times P}, \mathbf{B} \in \mathbb{R}^{J \times Q}, \mathbf{C} \in \mathbb{R}^{K \times R}$ は因子行列である．詳細は Kolda と Bader による解説 [Kolda 09] に詳しい．本実験では *TensorToolbox* [*4] を用いた．

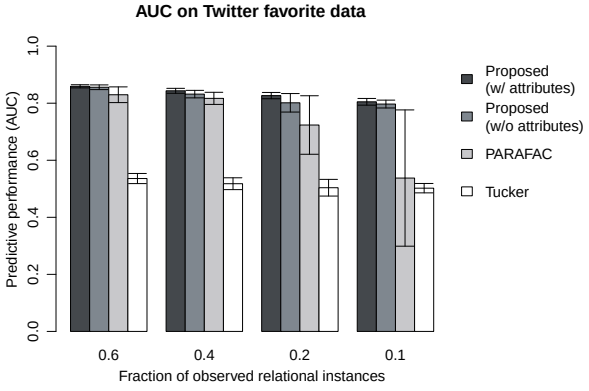
3·5 パラメータ設定

提案手法，比較手法それぞれについて，各サンプリング比率について，上記で定義した 1 スロットをチューニング用データとして用い，AUC の観点からパラメータを設定した．チューニングの際の結果は，実際の実験結果には含まれない．属性を活用しない提案手法では，R は $\{2^3, 2^4, \ldots, 2^9\}$ の 7 つを，属性を活用する提案手法では，

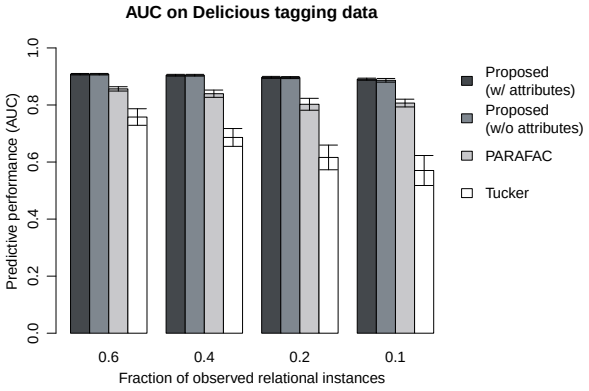
*2 http://norizm.org/datasets.html
*3 http://www.grouplens.org/node/462#attachments
*4 http://www.sandia.gov/~tgkolda/TensorToolbox/index-2.3.html



(a) Twitter における retweet アクション

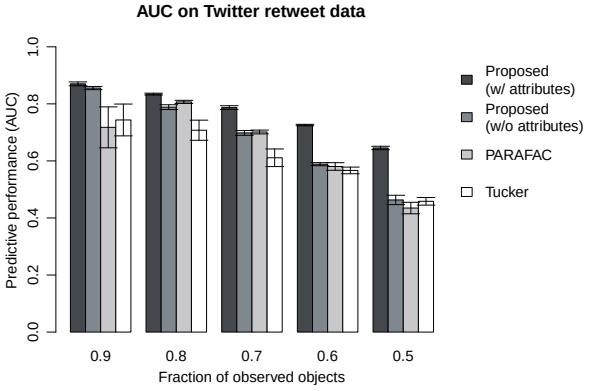


(b) Twitter における favorite アクション

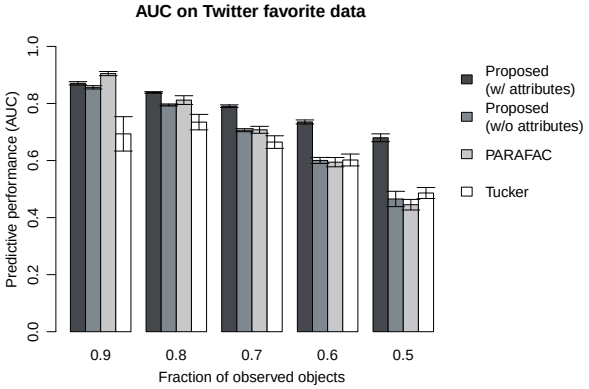


(c) Delicious における tagging アクション

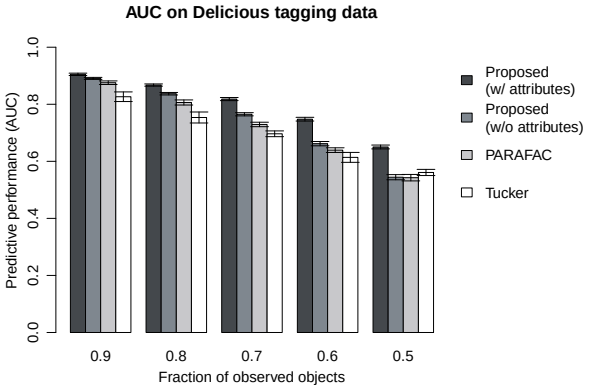
図 2 関係レベルで疎な状況での AUC の比較．属性の活用の有無に関わらず，提案手法がデータ疎性に対して最も高い頑強性を示した．



(a) Twitter における retweet アクション



(b) Twitter における favorite アクション



(c) Delicious における tagging アクション

図 3 オブジェクトレベルで疎な状況での AUC の比較．属性を活用した提案手法がデータ疎性に対して最も高い頑強性を示した．

R は $\{2^6, 2^7, \ldots, 2^9\}$ の 4 つを候補とした．提案手法で属性を活用する場合の σ は $\{10^{-2}, 10^{-3}, 10^{-4}\}$ で評価したが，このパラメータの範囲において安定した精度が確認されたため，実際の実験では，10^{-3} で固定した．比較手法においては R は，$\{2^0, 2^1, \ldots, 2^4\}$ の 5 つを候補とした．

3·6 結　　果

提案手法は，関係レベルで疎な状況では，属性の活用の有無に関わらず比較手法よりもデータ過疎に対して頑強な性能を示した．オブジェクトレベルで疎な状況では属性の活用が効果を発揮し，属性を活用した提案手法は比較手法よりも頑強な性能を示した．

図 2 が関係レベルで疎な設定での実験結果である．観測された関係インスタンスの割合を $0.6, 0.4, 0.2, 0.1$ と変えたときの AUC の平均を，標準偏差付きで示してある．この割合が 0.6 である場合というのは，全データセットの中で 6 割の関係インスタンスが訓練時に観測された場合を指す．全データセットにおいて，属性の有無に関わらず提案手法がデータ過疎への最も高い頑強性を示していることが確認できる．提案手法が見せている相対的に小さい標準偏差は，大域解を保証する定式化によるものと考えられる．詳細な分析として，各手法の全ペアについて t 検定 ($p < 0.05$) を行い以下の結論を得た．(1) 提案手法において属性の有無による精度の有意差は必ずしもなかった．データセットやサンプリング比率によって，属性の有無による有意差があるものとないものがあった．(2) 観測された関係インスタンスの割合が 4 割以下の全ての場合（6 割以外の場合）には，属性の有無に関わらず提案手法の比較手法に対する精度向上は有意であった．(3) 観測された関係インスタンスの割合が 6 割の場合は，提案手法（属性ありなし両方）の精度が比較手法の精度を有意に上回るか，両者の間に有意差がないかのいずれかであった．(1) について，提案手法において属性の活用の有無による有意差が必ずしも見られない理由としては，オブジェクトが既知である場合には，属性情報よりも関係データそれ自身の情報が関係形成において支配的になることが考えられる．(2)，(3) について，提案手法が属性の有無に関わらず，比較手法に対してデータ過疎への高い頑強性を示したのは，大域解を保証する定式化によるものと考えられる．

図 3 がオブジェクトレベルで疎な設定での実験結果である．予測時に既知であるオブジェクトの割合を

$0.9, 0.8, 0.7, 0.6, 0.5$ と変えたときの AUC の平均を，標準偏差付きで示してある．この割合が 9 割である場合というのは，全オブジェクトの中で 1 割のオブジェクトが，予測時に新しく観測された場合を指す．提案手法のうち属性を活用していない方の手法は，比較手法と同様にデータ過疎に対して低い頑健性を示しているが，属性を活用した提案手法はデータ過疎に対して相対的に高い頑強性を示した．関係レベルで疎な設定と同様に，詳細な分析として，各手法の全ペアについて t 検定 ($p < 0.05$) を行い以下の結論を得た．(1) 全ての場合で，属性を活用した提案手法の精度が，活用しない提案手法の精度を有意に上回った．(2) 既知のオブジェクトが 8 割以下の全ての場合（9 割以外の場合），すなわち未知のオブジェクトが 2 割以上存在する場合には，属性を活用した提案手法の精度がテンソルを用いた比較手法の精度を有意に上回った．(3) 既知のオブジェクトが 9 割の場合には，提案手法の精度がテンソルを用いた比較手法を有意に上回るか，両者の間に有意差がないかのいずれかであった．属性を活用した提案手法が比較手法に対してデータ過疎への高い頑強性を示したのは，未知のオブジェクトが多数存在するような状況では，属性の活用が有効であったためと考えられる．この結果は，属性の活用がコールドスタート問題を解決するにあたって有用であることを示唆している．

4. 関 連 研 究

多項関係予測においては，テンソルを用いた手法がよく使用されている [Dunlavy 11, Koutra 12, Sun 06a, Kolda 05, Sun 06b, Symeonidis 08, Rendle 09, Rendle 10]．これら，テンソル補完問題等のテンソル分析タスクにおいては，対象となるテンソルに対して低ランク性を仮定することが多く，様々な低ランク分解モデルが，効率的なアルゴリズムと共に提案されてきている [Kolda 09]．しかしながら，多くの既存手法で保証されるのは局所解のみであり，その予測性能は対象アルゴリズムに与える初期値に大きく依存する．対照的に，提案手法では一般化固有値問題を一度解くだけで大域解を得ることが可能になる．このような性質と属性情報の統合により，提案手法はデータ過疎に対して頑強な予測性能を実現した．

5. お わ り に

本論文では，異種のオブジェクト間に生じる多項関係を予測するために，関係データと，オブジェクトの属性情報の両方を活用する手法を提案した．提案手法は，大域解を保証する定式化とオブジェクトの属性を活用する定式化により，標準的なテンソル分解と比較してデータ過疎への高い頑強性を示した．

◇ 参 考 文 献 ◇

[Bao 07] Bao, S., Xue, G., Wu, X., Yu, Y., Fei, B., and Su, Z.: Optimizing web search using social annotations, in *Proceedings of the 16th International Conference on World Wide Web*, pp. 501–510 (2007)

[Belkin 03] Belkin, M. and Niyogi, P.: Laplacian Eigenmaps for Dimensionality Reduction and Data Representation, *Neural Computation*, Vol. 15, No. 6, pp. 1373–1396 (2003)

[Cai 11] Cai, Y., Zhang, M., Luo, D., Ding, C., and Chakravarthy, S.: Low-order tensor decompositions for social tagging recommendation, in *Proceedings of the fourth ACM International Conference on*

表 1 実験で用いたデータセットの詳細

データセット	タプル数	オブジェクト	オブジェクト数	属性	属性数
Twitter	14,221	subjective user	1,144	ユーザの tweet に含まれるキーワード	4,896
		mentioned user	7,935	follow されているユーザ	2,586
(retweet)		URL	11,335	URL と共起した subjective user の tweet に含まれるキーワード	4,757
Twitter	22,755	subjective user	1,125	ユーザの tweet に含まれるキーワード	4,107
		mentioned user	10,049	follow されているユーザ	2,586
(favorite)		URL	18,244	URL と共起した subjective user の tweet に含まれるキーワード	4,107
Delicious	33,414	user	768	friend 関係にあるユーザ	1,098
		tag	8,280	URL と共起したタグ	15,088
(tagging)		URL	6,860	URL と共起したユーザ	1,185

Web Search and Data Mining, pp. 695–704 (2011)
[Csanky 76] Csanky, L.: Fast Parallel Matrix Inversion Algorithms, *SIAM Journal on Computing*, Vol. 5, No. 4, pp. 618–623 (1976)
[Dunlavy 11] Dunlavy, D. M., Kolda, T. G., and Acar, E.: Temporal Link Prediction Using Matrix and Tensor Factorizations, *ACM Transactions on Knowledge Discovery from Data*, Vol. 5, No. 2, pp. 10:1–10:27 (2011)
[Guan 09] Guan, Z., Bu, J., Mei, Q., Chen, C., and Wang, C.: Personalized tag recommendation using graph-based ranking on multi-type interrelated objects, in *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 540–547 (2009)
[Heymann 08] Heymann, P., Koutrika, G., and Garcia-Molina, H.: Can social bookmarking improve web search?, in *Proceedings of the first International Conference on Web Search and Web Data Mining*, pp. 195–206 (2008)
[Kolda 05] Kolda, T. G., Bader, B. W., and Kenny, J. P.: Higher-order web link analysis using multilinear algebra, in *Proceedings of the fifth IEEE International Conference on Data Mining*, pp. 242–249, IEEE Computer Society (2005)
[Kolda 09] Kolda, T. G. and Bader, B. W.: Tensor Decompositions and Applications, *SIAM Review*, Vol. 51, No. 3, pp. 455–500 (2009)
[Koutra 12] Koutra, D., Papalexakis, E. E., and Faloutsos, C.: TensorSplat: Spotting Latent Anomalies in Time, in *Proceedings of the 16th Panhellenic Conference on Informatics*, pp. 144–149 (2012)
[Lehoucq 96] Lehoucq, R. and Sorensen, D. C.: Deflation Techniques For An Implicitly Re-Started Arnoldi Iteration, *SIAM Journal on Matrix Analysis and Applications*, Vol. 17, pp. 789–821 (1996)
[Mika 05] Mika, P.: Ontologies Are Us: A Unified Model of Social Networks and Semantics, in *Proceedings of the fourth International Semantic Web Conference*, Vol. 3729, pp. 522–536, Springer (2005)
[Pan 87] Pan, V. Y.: Complexity of Parallel Matrix Computations, *Theoretical Computer Science*, Vol. 54, pp. 65–85 (1987)
[Rendle 09] Rendle, S., Balby Marinho, L., Nanopoulos, A., and Schmidt-Thieme, L.: Learning optimal ranking with tensor factorization for tag recommendation, in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 727–736 (2009)
[Rendle 10] Rendle, S. and Schmidt-Thieme, L.: Pairwise interaction tensor factorization for personalized tag recommendation, in *Proceedings of the third ACM International Conference on Web Search and Data Mining*, pp. 81–90 (2010)
[Schein 02] Schein, A. I., Popescul, A., Popescul, R., Ungar, L. H., and Pennock, D. M.: Methods and Metrics for Cold-Start Recommendations, in *Proceedings of the 25th annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 253–260 (2002)
[Schifanella 10] Schifanella, R., Barrat, A., Cattuto, C., Markines, B., and Menczer, F.: Folks in Folksonomies: social link prediction from shared metadata, in *Proceedings of the third ACM International Conference on Web Search and Data Mining*, pp. 271–280 (2010)
[Sorensen 92] Sorensen, D. C.: Implicit application of polynomial filters in a k-step Arnoldi method, *SIAM Journal on Matrix Analysis and Applications*, Vol. 13, No. 1, pp. 357–385 (1992)
[Sun 06a] Sun, J., Papadimitriou, S., and Yu, P. S.: Window-based Tensor Analysis on High-dimensional and Multi-aspect Streams, in *Proceedings of the sixth IEEE International Conference on Data Mining*, pp. 1076–1080 (2006)
[Sun 06b] Sun, J., Tao, D., and Faloutsos, C.: Beyond streams and graphs: dynamic tensor analysis, in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 374–383 (2006)
[Symeonidis 08] Symeonidis, P., Nanopoulos, A., and Manolopoulos, Y.: Tag recommendations based on tensor dimensionality reduction, in *Proceedings of the 2nd ACM International Conference on Recommender Systems*, pp. 43–50 (2008)
[Voloshin 09] Voloshin, V. I.: *Introduction to Graph and Hypergraph Theory*, Nova Kroshka Books (2009)
[White 05] White, S. and Smyth, P.: A spectral clustering approach to finding communities in graphs, in *Proceedings of the fifth SIAM International Conference on Data Mining* (2005)
[Yamanishi 09] Yamanishi, Y.: Supervised Bipartite Graph Inference, in *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems*, pp. 1841–1848 (2009)

〔担当委員：小町 守〕

2013 年 7 月 23 日 受理

著 者 紹 介

則 のぞみ（学生会員）

2010 年東京大学工学部システム創成学科卒．2012 年東京大学大学院情報理工学系研究科創造情報学専攻修士課程終了．広く情報学と社会学の接点に興味を持つ．現在はグラフ/ネットワーク解析，Web マイニング，機械学習に関する研究を行っている．ACM 会員．

ボレガラ ダヌシカ（正会員）

2005 年東京大学工学部電子情報工学科卒．2007 年同大学院情報理工学系研究科修士課程修了．2009 年同研究科博士課程修了．博士 (情報理工学)．同研究科・助教，講師を経て現在は University of Liverpool (Department of Computer Science) の Senior Lecturer (Associate Professor)．自然言語処理に興味を持つ．WWW, ACL, ECAI などの会議を中心に研究成果を発表．

鹿島 久嗣（正会員）

1999 年京都大学大学院工学研究科応用システム科学専攻修士課程修了．2007 年京都大学大学院情報学研究科知能情報学専攻博士課程修了．博士 (情報学)．1999 年から 2009 年まで日本アイ・ビー・エム株式会社 東京基礎研究所勤務．2009 年より東京大学大学院情報理工学系研究科数理情報学専攻・准教授．機械学習やデータマイニングの研究，特にグラフやネットワーク構造をもったデータを対象とする予測モデリングに取り組む．