

図書館の貸出履歴を用いた 図書推薦システムの有効性検証

辻慶太, 黒尾恵梨香, 佐藤翔, 池内有為, 池内淳, 芳鐘冬樹, 逸村裕

Effectiveness of using library loan records for book recommendation, by TSUJI Keita, KUROO Erika, SATO Sho, IKEUCHI Ui, IKEUCHI Atsushi, YOSHIKANE Fuyuki, and ITSUMURA Hiroshi.

本研究では(1)T大学附属図書館の貸出履歴約5年分を用いた協調フィルタリング, (2)同データを用いたアソシエーションルール, (3)Amazon, の3つによって被験者33名に図書を推薦し, それぞれの有効性を比較検証した。結果, (3) > (2) > (1)の順に推薦パフォーマンスが高いことを確認した。貸出履歴を用いて図書推薦を行うならば, 協調フィルタリングよりも, プライバシー漏洩の可能性が低く計算コストも少ないアソシエーションルールを用いた方が有効であること, さらにAmazonの利用・併用も検討に値することが言えた。

1. はじめに

近年図書館界では, 利用者の貸出履歴を有効に利用する方法が活発に議論されている。¹⁾²⁾³⁾⁴⁾⁵⁾ 其中之一つに貸出履歴を用いた図書推薦システムの実現がある。ここで貸出履歴とは, 各利用者が借りた資料のID, 書誌事項, 借りた年月日から構成される利用者ごとの貸出記録の総体を指し, 図書推薦とは図書のタイトル・著者名・出版者名・刊年といった書誌事項を提示して, その図書の閲覧・貸出を利用者に薦めることを指す。実際に貸出履歴を用いて図書推薦を行っている図書館も存在し, 推薦手法を提案する研究も存在する。⁶⁾⁷⁾⁸⁾ だが, (1)貸出履歴をどのように用いて推薦を行うのが有効か, (2)そもそも貸出履歴を用いた推薦は他の手法による推薦より有効なのかといった検証はあまり行われていない。以下ではこれら2点について述べる。

まず(1)だが, 一般に何らかのアイテム(商品など)の利用・購入履歴を用いて顧客に推薦を行う手法には, 協調フィルタリングやアソシエーションルールなど様々な手法がある。ここで協調フィルタリングとは, (i)顧客が利用・購入したアイテムを顧客ごとに記録しておき, (ii)推薦対象者と過去の利用・購入アイテムが似ている顧客群を特定し, (iii)その顧客達が利用・購入し, かつ推薦対象者が利用・購入していないアイテムを推薦すること, を指す。⁹⁾ またアソシエーションルールとは, 顧客が同時に利用・購入したアイテム集合を記録しておき, あるアイテム(集合)を利用・購入しようとする推薦対象者に対して, 過去に同時に利用・購入されるケースが多かったアイテム(集合)を推薦することを指す。アソシエーションルールはどの顧客がそのアイテム集合を利用・購入したかという個人情報を保持しておく必要がない。

さて, だが図書の貸出履歴に関してこうした推薦手法のいずれが最も有効なのかという検証は行われていない。例えば原田・増田は貸出履歴を用いた協調フィルタリングで図書を推薦する手法を提案しているが, 他の手法との比較は行っていない。¹⁰⁾ 各手法には推薦パフォーマンスの違いの他にも様々な長所短所がある。例えば協調フィルタリングは, 個々人の貸出履歴を保持し続ける必要があり, プライバ

2012年6月28日受理

つじ けいた 筑波大学図書館情報メディア系
くろお えりか 筑波大学情報学群知識情報・図書館学類
さとう しょう 筑波大学大学院図書館情報メディア研究科
いけうち うい 筑波大学大学院図書館情報メディア研究科
いけうち あつし 筑波大学図書館情報メディア系
よしかね ふゆき 筑波大学図書館情報メディア系
いつむら ひろし 筑波大学図書館情報メディア系

September 2012

シー漏洩の危険が常に伴う短所を持つ。一方アソシエーションルールは、一緒に貸し出された図書の組み合わせ情報だけで実現でき、個々人の貸出履歴を保持し続ける必要がない。また協調フィルタリングはアソシエーションルールに比べて計算量が多く、推薦図書の決定に時間がかかる。以上のことから、もしアソシエーションルールの方が協調フィルタリングよりも推薦パフォーマンスが高かったならば、図書館はアソシエーションルールの方を用いた方が効果的かつ安全・高速に推薦を行えると言える。

次に(2)について述べる。図書館外部の図書推薦システムとしては Amazon のそれが有名である。¹¹⁾¹²⁾ Amazon は広く浸透したサービスであり、様々な情報を利用できる API も公開している。図書館の蔵書のうち利用者が興味を持ったもの(例えば借りたり OPAC でクリックしたもの)に対して、Amazon が「その図書を購入した人は以下の図書も購入しています」と提示する図書を API で調べ、それらの図書の中から図書館が所蔵するものを表示すれば、ある種の図書推薦システムを実現することが出来る。もしそのような推薦の方が貸出履歴を用いた推薦より利用者評価が高かったならば、(Amazon が消滅したり情報提供をやめたり、推薦方法を内部で大きく変更したりしない限り) 図書館は貸出履歴を用いることなく効果的に図書の推薦を行えると言える。

以上の背景・問題関心から本研究では、後述する T 大学の学生33名を被験者として、彼らに現在興味を持っている図書を1冊ずつ挙げてもらい、その図書と彼らの T 大学附属図書館における貸出履歴を用いて図書推薦の実験を行った。実験では(a)原田・増田が提案する協調フィルタリング、¹³⁾ (b)アソシエーションルール、(c)Amazon、の3つを用いて推薦を行い、被験者にその適切さを評価してもらった。もし(b)(c)による推薦の方が(a)による推薦よりも評価が高かったならば、本研究は図書推薦システムのあり方に関し先述のような知見をもたらす。即ち、貸出履歴を用いてより安全・高速・効果的に推薦するシステムや、そもそも貸出履歴を用いず効果的に推薦するシステムのあり方に関する知見である。なお本研究では館種としては大学図書館を取り上げ実験調査の対象とする。ここで得られた知見のいくつかは公共図書館においても有効と思われるが、その精査・検証は今後の課題としたい。

辻ほか：図書館の貸出履歴を用いた図書推薦システムの有効性検証

本論文は以下のように構成される。まず次の第2章では先行研究を紹介する。第3～5章では本研究に用いたデータや推薦手法及び評価手法について述べ、第6章で結果を述べる。第7章では結果を踏まえた考察を行い、第8章で総括する。

2. 先行研究

協調フィルタリングを紹介する文献では Resnick らの手法がしばしばその代表例として取り上げられる。¹⁴⁾¹⁵⁾ Resnick らは、Netnews の読者が各記事に付与した評価値を用い、まず似たような評価を行った読者同士を嗜好が類似した読者とみなす。そしてある読者に対し、類似読者が高い評価値を付与している記事を勧めることを提案している。こうした手法の詳細については次章で述べる。

Resnick ら以降、様々な分野で協調フィルタリングの利用研究が行われてきたが、図書館における図書推薦に協調フィルタリングを用いる研究は少ない。わずかに先述の原田、¹⁶⁾ 原田・増田、¹⁷⁾ 及び辻ら¹⁸⁾の研究がある程度である。これは貸出履歴の利用が図書館界では禁忌に近かったことも影響しているかもしれない。以下ではこれら3つの研究について述べる。

まず原田は貸出履歴として、ある大学図書館の利用者8,808人分の960,078件のデータを用いている。¹⁹⁾ 被験者は「公開できる貸出図書数が21冊以上ある」大学生35名で、彼らにそれぞれ30冊の図書を推薦し評価してもらっている。推薦図書の決定は Vogoo PHP Pro v2.2という協調フィルタリングソフトで行っている。このソフトと貸出冊数21冊以上という条件については後述する。

次に原田・増田について述べる。²⁰⁾ 原田・増田も Vogoo PHP Pro v2.2を用いて協調フィルタリングによる推薦を行っている。貸出履歴としては、ある大学図書館でこれまで200冊以上500冊未満の図書を借りた学生の貸出履歴272,751件を用いている。被験者は先ほど同様公開できる貸出図書数が21冊以上ある大学生10名である。先ほどの原田と大きく異なる点は、原田・増田は図書の NDC(日本十進分類法)の類を重み付けに用い、借りた年月によって貸出履歴を分割している点である(この点については後述する)。原田・増田はこうした重み付けや分割に関して5種類のパターンを試し、最も評価が高いパターンを特定している。だが先ほどの原田同様、Amazon やアソシエーションルールとの比較など

は行っていない。

次らは本研究と同じ貸出履歴と協調フィルタリングを用い、Amazon の推薦との比較を行っている。²¹⁾ だが被験者の数は4名と少なく、またアソシエーションルールとの比較も行っていない。

さて協調フィルタリングを用いずに図書館での図書推薦を行う研究としては Whitney & Schiff があ²²⁾る。Whitney & Schiff は UCLA (University of California, Los Angeles) 附属図書館の貸出トランザクションをデータとして用い、手法としてはグラフモデルに基づく推薦手法を提案している。彼らの手法は本研究が取り上げるアソシエーションルールと本質的に近い。

実験を行わずに図書推薦の手法を提案した文献としては、Shirgaonkar ら²³⁾、Chen & Chen²⁴⁾、Luo ら²⁵⁾などがある。また協調フィルタリングなど推薦アルゴリズムの解説としては神嶌が詳しい。^{26) 27) 28)}

3. 用いたデータ

3.1 貸出履歴

本研究では、T 大学附属図書館の貸出記録をもとに図書の推薦実験を行う。²⁹⁾ 1 件の貸出記録は資料 1 タイトルの書誌情報とある年月日にその資料を借りた利用者 1 名の情報から成る。具体的にはその資料の資料番号、著者名、タイトル、請求番号といった情報と、その図書を借りた利用者の利用者番号、利用者の所属、学年、貸出年月日といった情報から成る。上記貸出記録は2006年1月2日から2011年3月31日までのものであり、全部で1,990,797件あった。このうち学部生によるものは864,704件、大学院生・教職員によるものは989,641件、その他利用者(卒業生・一般学者・名誉教授など)によるものは136,452件であった。本研究では「その他利用者」による貸出記録は用いず、学部生・大学院生・教職員による貸出記録1,854,345件(=864,704件+989,641件)を用いることにした。³⁰⁾ これらの異なり図書数は435,817冊、異なり利用者数は39,442人(うち学部生は19,077人)であった。後述するが、上記のうち協調フィルタリングの参照情報源となった「200冊以上500冊未満の図書を借りた者」の数は1,678人であり、彼らが借りた図書の延べ数は488,565冊であった。

T 大学附属図書館の貸出記録は先述のような形式を取っていることから、1 人の利用者が同時に借りた資料の組(いわゆる 1 バスケット中のアイテム)

は厳密には分らない。そこで本研究では同じ利用者が同じ年月日に借りた資料群は同時に借りられたとみなすことにした。学部生・大学院生・教職員によるバスケット数は708,951であった。

3.2 現在借りたい 1 冊

先述のように本研究では協調フィルタリング、アソシエーションルール、Amazon による図書推薦の有効性を比較する。後述するが有効性の検証は33名の被験者に図書を推薦し、その適切さを評価してもらう形で行った。各被験者には研究や学習で用いる為に現在 T 大学附属図書館で借りたい図書を 1 冊挙げてもらった。これも後述するが、アソシエーションルールと Amazon による推薦ではこの 1 冊を主要な手がかりとした。³¹⁾

4. 推薦方法

4.1 協調フィルタリングを用いた原田・増田の手法による推薦

以下では本研究で採用した協調フィルタリング手法について述べる。本研究では、原田・増田の手法を、貸出履歴に基づく協調フィルタリングを用いた推薦手法の代表例として取り上げたい。³²⁾ だが原田・増田は協調フィルタリングに Vogoo PHP Pro v2.2 という市販ソフトを用いたと述べているだけで、計算方法の詳細は記していない。同ソフトは現在販売会社が消滅しているため入手が難しい。そこで本研究では、Resnick らの手法を協調フィルタリングの計算方法に採用し、評価の重み付けに原田・増田の手法を用いることで彼らのシステムの再現とした。³³⁾ Resnick らの手法は、先述の神嶌の解説論文でも紹介されている協調フィルタリングの代表的手法である。³⁴⁾ 以下ではこの Resnick らの手法について説明し、次に原田・増田の重み付け手法について説明する。

4.1.1 Resnick らの手法

以下では推薦対象利用者を活動利用者と呼び、他の利用者を標本利用者と呼ぶ。Resnick らの手法は、基本的に活動利用者と嗜好の類似度が高い標本利用者が高い評価を下し、かつ活動利用者がこれまで評価を下していないアイテムを推薦する。以下ではその計算方法と簡単な計算例を示す。

まず以下のように記号を定義する：

Y_{ai} ：活動利用者 a と標本利用者 i が共通して評

September 2012

価したアイテムの集合。

S_{aj} , S_{ij} :それぞれ活動利用者 a , 標本利用者 i がアイテム j に対して付与した評価値。

$\bar{S}_a$, $\bar{S}_i$:それぞれ活動利用者 a , 標本利用者 i が付与した全評価値の平均。

X_j :アイテム j を評価した利用者の集合。

活動利用者 a と標本利用者 i の類似度 p_{ai} は、両者が共に評価したアイテムに関する以下の相関係数とする。

$$p_{ai} = \frac{\sum_{k \in Y_{ai}} (S_{ak} - \bar{S}_a)(S_{ik} - \bar{S}_i)}{\sqrt{\sum_{k \in Y_{ai}} (S_{ak} - \bar{S}_a)^2} \sqrt{\sum_{k \in Y_{ai}} (S_{ik} - \bar{S}_i)^2}} \quad \dots (1)$$

活動利用者がこれまで評価していないアイテム j に対する推定評価値 $\hat{S}_{aj}$ は以下のように定義する。Resnick らの手法はこの値が高いアイテムを活動利用者に推薦する。

$$\hat{S}_{aj} = \bar{S}_a + \frac{\sum_{i \in X_j} p_{ai} (S_{ij} - \bar{S}_i)}{\sum_{i \in X_j} |p_{ai}|} \quad \dots (2)$$

表 1：評価値の例

	図書 A	図書 B	図書 C	図書 D
利用者 1	1	3	*	3
利用者 2	*	1	3	*
利用者 3	2	1	3	1
利用者 4	1	3	2	*

以下に簡単な計算例を示す。まず表 1 は利用者 1 ～ 4 が図書 A～D に対して付与した評価値である。評価値は {1, 2, 3} の 3 段階とし、* は未評価であることを表す。表 1 から例えば利用者 2 は図書 C に対し評価値 3 という高い値を付与したものの、図書 B には 1 という低い値を付与したこと、図書 A, D はまだ評価していないことが分かる。

ここで活動利用者、即ち推薦対象者を利用者 2 とし、利用者 1, 3, 4 を標本利用者、即ち利用者 2 に対して推薦する図書を決める為のサンプルとする。この時、利用者 2 に対する図書 A の推定評価値 $\hat{S}_{2,A}$ は以下のように算出する。まず図書 A を既に評価している標本利用者 1, 3, 4 と活動利用者 2 との類似度を式 (1) により算出する。標本利用者 1 と活動利用者 2 とでは共通に評価している図書が B のみの 1 冊であるため $p_{2,1} = 0$ とする。標本利用者 3 と活動利用者 2 とでは共通に評価している図書

は C: 図書の貸出履歴を用いた図書推薦システムの有効性検証

が B, C であり、 $Y_{2,3} = \{B, C\}$ となる。これらの図書に関する利用者 2, 3 それぞれの平均評価値 $\bar{S}_2$, $\bar{S}_3$ は以下ようになる:

$$\bar{S}_2 = (\sum_{k=B,C} S_{2,k}) / 2 = (1 + 3) / 2 = 2$$

$$\bar{S}_3 = (\sum_{k=B,C} S_{3,k}) / 2 = (1 + 3) / 2 = 2$$

よって利用者 2, 3 の類似度は以下ようになる:

$$\begin{aligned} p_{2,3} &= \frac{\sum_{k=B,C} (S_{2,k} - \bar{S}_2)(S_{3,k} - \bar{S}_3)}{\sqrt{\sum_{k=B,C} (S_{2,k} - \bar{S}_2)^2} \sqrt{\sum_{k=B,C} (S_{3,k} - \bar{S}_3)^2}} \\ &= \frac{(1-2)(1-2) + (3-2)(3-2)}{\sqrt{(1-2)^2 + (3-2)^2} \sqrt{(1-2)^2 + (3-2)^2}} \\ &= 1 \end{aligned}$$

同様に標本利用者 4 と活動利用者 2 の類似度 $p_{2,4}$ は -1 となる。これらを式 (2) に代入すると活動利用者 2 の図書 A に対する推定評価値 $\hat{S}_{2,A}$ は以下ようになる:

$$\begin{aligned} \hat{S}_{2,A} &= \bar{S}_2 + \frac{\sum_{i=1,3,4} p_{2,i} (S_{i,A} - \bar{S}_i)}{\sum_{i=1,3,4} |p_{2,i}|} \\ &= 2 + \frac{0(1-3) + 1(2-2) + (-1)(1-5/2)}{|0| + |1| + |-1|} \\ &= 2.75 \end{aligned}$$

この値は評価値の最大値 3 に近いことから、利用者 2 は図書 A を好む可能性が高いなどと判断することが出来る。

4.1.2 原田・増田の重み付け手法

本研究では上記評価値として原田・増田が言うところの「パターン D」の重みを採用した。³⁵⁾「パターン D」は、(1)図書の分野と、(2)標本利用者の経時の変化を踏まえた重みであり、彼らの実験では最も効果的な推薦を行っていた。以下ではこの重み付け手法について述べる。

まず(1)の図書の分野については、活動利用者が借りた図書の NDC に関して k 番目に多い類の図書に $0.5^{(k-1)}$ の重みを付与する。即ち、最も多い類の図書には 1.0 を付与し、2 番目 3 番目に多い類の図書にはそれぞれ 0.5, 0.25 を付与する。次に(2)の経時の変化については 200 冊以上 500 冊未満の図書を借りた者を標本利用者とし、さらに貸出履歴を一月単位で分割し、異なる年月の貸出履歴は異なる標本利用者の貸出履歴とみなす。例えば同一の標本利用者であっても 2009 年 6 月、2009 年 11 月、2010 年 6 月の貸出履歴はそれぞれ異なる 3 つの標本利用者の貸

出履歴とみなす。標本利用者が借りた図書には評価値1を付与する。

さて原田・増田は上記のように、標本利用者が借りた図書には評価値1を付与しているが、借りていない図書にどのような評価値を与えたかは明記していない。上記のように、標本利用者が借りた図書に評価値として一律1を付与した場合、前節の類似度 p_{ai} は分母が0になり算出することが出来ない。そこで本研究では借りていない図書に評価値0を便宜上付与することにした。この妥当性については今後の課題としたい。

本研究では以上の手法によって算出された推定評価値が、最も高い6冊を各被験者に提示した。

4.2 アソシエーションルールによる推薦

以下では本研究が採用したアソシエーションルールによる推薦方法について述べる。³⁶⁾³⁷⁾まず n 冊の図書 $X_i (i=1, \dots, n)$ をある利用者が一緒に借りた場合、その貸出履歴を $\{X_1, \dots, X_n\}$ と表し、これをトランザクションと呼ぶ。例えば図書 A, B, C を一緒に借りた場合のトランザクションは $\{A, B, C\}$ と表す。さてこのトランザクションからは「図書 A を借りる利用者は図書 B も借りる」というルールを引き出すことが出来る(その確かさについては後述する)。また「図書 B, C を同時に借りる利用者は図書 A も借りる」というルールを引き出すことも出来る。このようにある図書集合 X を借りる利用者は図書集合 Y も借りる(ただし $X \cap Y = \emptyset$ かつ、 $X \neq \emptyset, Y \neq \emptyset$)というルールを $X \rightarrow Y$ と表すと、上記トランザクションからは $\{A\} \rightarrow \{B\}, \{A\} \rightarrow \{C\}, \{B\} \rightarrow \{C\}, \{C\} \rightarrow \{B\}, \{C\} \rightarrow \{A\}, \{B\} \rightarrow \{A\}, \{A\} \rightarrow \{B, C\}, \{B\} \rightarrow \{A, C\}, \{C\} \rightarrow \{A, B\}, \{A, B\} \rightarrow \{C\}, \{A, C\} \rightarrow \{B\}, \{B, C\} \rightarrow \{A\}$ という12個のルールを引き出すことが出来る。矢印の左側を前提部、右側を結論部と呼ぶ。

さてルール $X \rightarrow Y$ の確かさ、有益さについては確信度、支持度、リフトの3つで評価することが出来る。それらは以下のように定義される：

$$\text{確信度} = \frac{X, Y \text{ を構成する図書を含んだトランザクションの数}}{X \text{ を構成する図書を含んだトランザクションの数}}$$

$$\text{支持度} = \frac{X, Y \text{ を構成する図書を含んだトランザクションの数}}{\text{全トランザクションの数}}$$

$$\text{リフト} = \text{確信度} \times \frac{\text{全トランザクション数}}{Y \text{ を構成する図書を含んだトランザクションの数}}$$

例えば $\{A, B, C\}, \{A, B\}, \{B, C\}, \{A, C, D\}$ という4つのトランザクションがあった場合、 $\{A\} \rightarrow \{B\}$ の確信度は $2/3=0.67$ 、支持度は $2/4=0.50$ 、リフトは $(2/3) \times (4/3)=0.89$ となる。確信度が高ければ X を借りた利用者の多くは Y も借りていたことになり、従って今後 X を借りた利用者(あるいは興味を持っている利用者)に Y を示せば、その推薦は有益と判断される可能性が高い。

支持度が高いルールは多くの利用者がその図書の組合せを借りたということである。従って $X \rightarrow Y$ と $X \rightarrow Y'$ の確信度が同じで前者の方が支持度が高かったならば、 X を借りた利用者(あるいは興味を持っている利用者)には Y' よりも Y を推薦した方が有益と判断される可能性が高い。だが例えば全てのトランザクションに Y が含まれていたならば、 $X \rightarrow Y$ の確信度と支持度が共に高かったとしても X に対して Y を殊更推薦する意味はない。リフトは確信度と支持度の高さに意味があるかを判定する尺度である。リフトの値が1に近ければあまり有益なルールではなく、値が高いならば有益なルールとみなすことが出来る。

本研究ではまず、ルールの前提部と結論部が共に1冊の図書で、確信度が0.01以上、支持度が0.000001以上、リフトが1以上のルールを抽出し、³⁸⁾それらを確信度>支持度>リフトの優先順位で降順に並べた。次に前提部が被験者の「現在借りたい1冊」になっているルールを上から6つ抽出し、³⁹⁾それらの結論部の図書を被験者に推薦した。

4.3 Amazon による推薦

被験者の「現在借りたい1冊」をAmazonで検索し、「この商品を買った人はこんな商品も買っています」と表示された6冊のうちT大学附属図書館が所蔵しているものをOPACで特定し、各被験者に推薦した。⁴⁰⁾上記の検索及び特定は、本研究では手作業で行ったが、Amazonが提供しているAPIを用いれば自動で行えるようになるはずである。⁴¹⁾

5. 評価方法

以下ではまず本研究の被験者について述べ、次に被験者への推薦図書の提示方法と判定カテゴリーに

September 2012

ついて述べる。

5.1 被験者

本研究では T 大学で図書館情報学を専攻している学生33名に図書を推薦し、その適切さを判定してもらった。これら被験者33名の内訳は大学院生7名、学部4年生17名、学部2年生9名である(なお以下ではこれら3種類のグループは便宜上「学年」が異なると呼ぶことにする)。原田・増田は被験者を「過去に21冊以上図書館で図書を借りた学生」に限定している⁴²⁾。だが本研究では過去に借りた冊数と推薦の適切さの関係も調べるため、そのような制限は設けなかった。

今回の被験者が過去に借りた冊数は表2のようにまとめられた。表2から分かるように33名の被験者中13名(39.4%)は過去に20冊も借りていない。また貸出記録に現れた学部生19,077名のうち、20冊以下しか借りていない者は9,538名で50.0%を占めた。もし原田・増田の手法が、過去に21冊以上借りた学生にしか適切な推薦が行えないならば、彼らの手法は推薦できる学生が比較的限られていることになる。この点については後述する。

5.2 被験者への提示方法と判定カテゴリー

被験者には、先述の協調フィルタリング、アソシエーションルール、Amazon それぞれによって推薦された図書の書誌事項を提示した。提示書誌事項は原田・増田と同様、図書のタイトル、著者、出版者、刊年の4つである。各図書が先述の3手法のいずれによって推薦されたかは被験者には分からない(そもそも推薦に3つの手法を用いたことも被験者には知らせていない)。

表2：被験者の学年および貸出冊数

学年	過去に借りた冊数															
院 生	238	207	108	67	58	58	46	17								
4 年 生	361	82	68	64	60	45	37	27	22	22	11	9	8	6	1	1
2 年 生	82	76	31	10	4	3	1	1	0							

表3：全体の結果 カッコ内は%(以下、表4～18も同じ)

	△：既読	2：強い興味	1：読みたい	0：興味なし	×：評価不能	合計
Amazon	13(10.7)	35(28.7)	49(40.2)	21(17.2)	4(3.3)	122
アソシエーション	6(4.3)	26(18.6)	52(37.1)	52(37.1)	4(2.9)	140
協調フィルタリング	11(5.6)	22(11.1)	52(26.3)	94(47.5)	19(9.6)	198

辻ほか：図書館の貸出履歴を用いた図書推薦システムの有効性検証

被験者には、提示された書誌事項に基づいて、各図書の推薦の適切さを以下の5つのカテゴリーで判定してもらった：

2：「強い興味」

1：「読みたい」

0：「興味なし」

×：内容の判断ができないなどの理由で「評価不能」

△：既に関して持っている、または読んだことがある

これらは原田・増田のものとはほぼ同じである。異なるのは△の「既に関して持っている、または読んだことがある」というカテゴリーを追加した点である。現在の興味の強弱に関わらず、既に関して持っている、または読んだことがある図書の場合は△を選んでもらった。

6. 結果

以下ではまず全体の結果を示した後、学年別、これまで借りた冊数別、協調フィルタリングの推定評価値別、アソシエーションルールの確信度・支持度別、分野別の結果を示す。

6.1 全体の結果

Amazon、アソシエーションルール、協調フィルタリングによる推薦の評価結果は表3のようになった。

「×：評価不能」とされた図書は、タイトルや著者名だけでは少なくとも被験者の興味を惹かなかった図書であり、また洋書が多く、被験者が挙げた「現

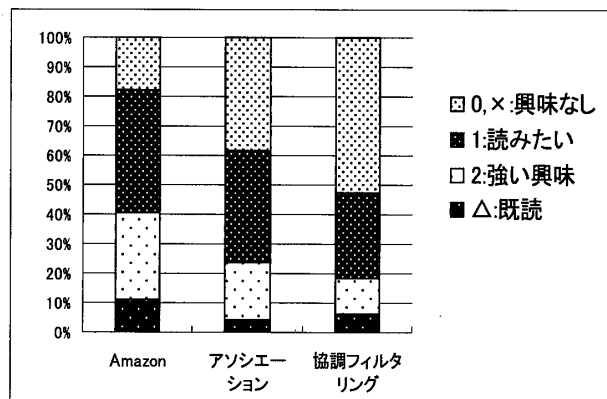


図1：全体の結果

在借りたい1冊」とは主題が離れている場合が多かった。そこで「0：興味なし」と合わせて図示すると、表3の結果は図1のようになった。

表3および図1に示すように、「2：強い興味」という評価を受けた図書の割合はAmazon, アソシエーションルール, 協調フィルタリングでそれぞれ28.7%, 18.6%, 11.1%となり、この順に高い値となっている。同様に「1：読みたい」についてもそれぞれ40.2%, 37.1%, 26.3%となり、Amazon, アソシエーションルール, 協調フィルタリングの順に高い値となっている。「Δ：既読」「2：強い興味」「1：読みたい」の3つを肯定的評価とみなすと、その割合はAmazon, アソシエーションルール, 協調フィルタリングそれぞれ79.6%(=10.7+28.7

+40.2), 60.0%, 43.0%で、この順に高いことが分かる。これら比率の間には有意水準0.01で差が認められた。⁴³⁾以上のことから全般に、協調フィルタリングよりもアソシエーションルールの方が結果が良く、Amazonの方がさらに結果が良いと言える。

6.2 学年別の結果

被験者を大学院生、学部4年生、2年生に分けて集計したところ表4～表6及び図2～図4のような評価結果となった。これらの表と図から、先述の肯定的評価の割合は大学院生、学部4年生、2年生いずれにおいてもAmazon, アソシエーションルール, 協調フィルタリングの順に高いことが分かる。例えば学部4年生におけるこれらの肯定的評価の割合はそれぞれ76.7%(=11.7+21.7+43.3), 56.0%, 31.3%となっている。これらの間には有意水準0.01で差が認められた。

協調フィルタリングに関する肯定的評価の割合は大学院生、学部4年生、2年生それぞれ56.2%(=12.5+8.3+35.4), 31.3%, 51.9%となっている。即ち、学年に応じて同割合が単調に増加あるいは減少するという傾向は見られなかった。

6.3 貸出冊数別の結果

先述のように原田・増田は貸出冊数が21冊以上の学生に被験者を限定している。そこで本研究でも被

表4：大学院生に関する結果

	Δ：既読	2：強い興味	1：読みたい	0, X：興味なし	合計
Amazon	5(17.2)	9(31.0)	10(34.5)	5(17.2)	29
アソシエーション	4(11.1)	5(13.9)	14(38.9)	13(36.1)	36
協調フィルタリング	6(12.5)	4(8.3)	17(35.4)	21(43.8)	48

表5：学部4年生に関する結果

	Δ：既読	2：強い興味	1：読みたい	0, X：興味なし	合計
Amazon	7(11.7)	13(21.7)	26(43.3)	14(23.3)	60
アソシエーション	2(3.0)	7(10.6)	28(42.4)	29(43.9)	66
協調フィルタリング	3(3.1)	9(9.4)	18(18.8)	66(68.8)	96

表6：学部2年生に関する結果

	Δ：既読	2：強い興味	1：読みたい	0, X：興味なし	合計
Amazon	1(3.0)	13(39.4)	13(39.4)	6(18.2)	33
アソシエーション	0(0.0)	14(36.8)	10(26.3)	14(36.8)	38
協調フィルタリング	2(3.7)	9(16.7)	17(31.5)	26(48.1)	54

September 2012

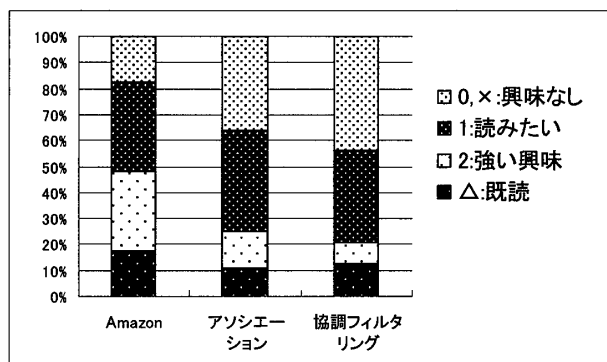


図2：大学院生に関する結果

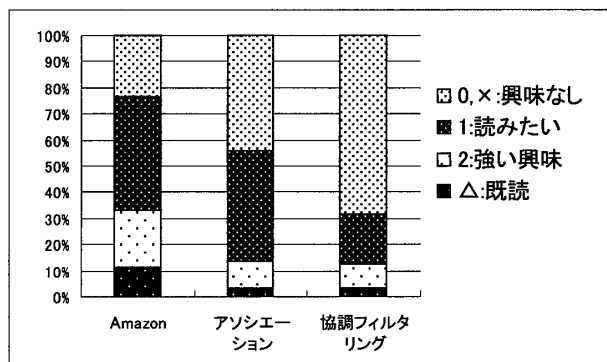


図3：学部4年生に関する結果

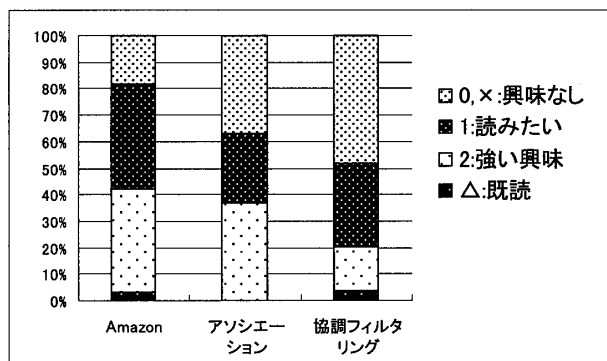


図4：学部2年生に関する結果

辻ほか：図書館の貸出履歴を用いた図書推薦システムの有効性検証

験者を貸出冊数が21冊以上のグループ(20名)と20冊以下のグループ(13名)の2つに分け、それぞれにおける評価結果をまとめた。結果はそれぞれ表7と図5、表8と図6のようになった。協調フィルタリングに関する肯定的評価の割合は貸出冊数が21冊以上のグループでは49.2%(=6.7+9.2+33.3)(表7)、貸出冊数が20冊以下のグループでは33.3%(表8)であることが分かる。これら比率の間には有意水準0.05で差が認められた。即ち、貸出冊数が多いグループに比べて、少ないグループの方では、協調フィルタリングのパフォーマンスが落ちることが示

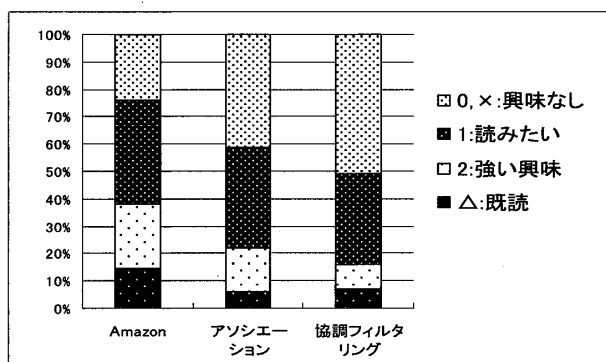


図5：貸出冊数が21冊以上のグループに関する結果

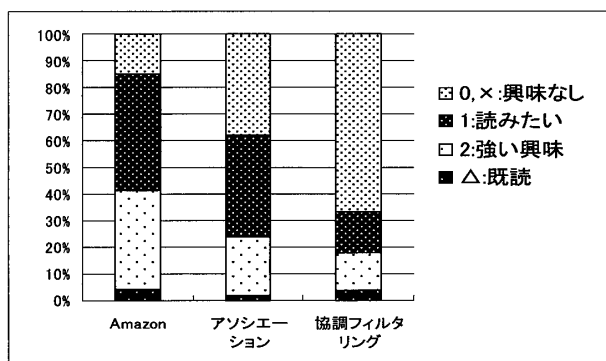


図6：貸出冊数が20冊以下のグループに関する結果

表7：貸出冊数が21冊以上のグループに関する結果

	Δ：既読	2：強い興味	1：読みたい	0，X：興味なし	合計
Amazon	11(14.5)	18(23.7)	29(38.2)	18(23.7)	76
アソシエーション	5(6.1)	13(15.9)	30(36.6)	34(41.5)	82
協調フィルタリング	8(6.7)	11(9.2)	40(33.3)	61(50.8)	120

表8：貸出冊数が20冊以下のグループに関する結果

	Δ：既読	2：強い興味	1：読みたい	0，X：興味なし	合計
Amazon	2(4.3)	17(37.0)	20(43.5)	7(15.2)	46
アソシエーション	1(1.7)	13(22.4)	22(37.9)	22(37.9)	58
協調フィルタリング	3(3.8)	11(14.1)	12(15.4)	52(66.7)	78

された。先述のように貸出冊数が20冊以下の学生は本研究の被験者の約4割を占め、貸出記録に現れた学部生の約5割を占める。本研究の被験者や学部生が日本の大学生の代表的なサンプルであった場合、日本の学生に対する協調フィルタリングの推薦結果は原田・増田が示したもののより低い可能性がある。

6.4 協調フィルタリングにおける推定評価値別の結果

協調フィルタリングでは、被験者は推定評価値が高い図書ほど好むと考えられる。本研究で協調フィルタリングが推薦した図書の推定評価値は、平均0.06、最大値0.26、最小値0.000063、中央値0.044であった。そこで推薦する図書を推定評価値が0.1以上のものに限った場合と、0.05以上に限った場合について評価結果をまとめた。結果は表9、図7のようになった。表9から、推定評価値が0.1以上の図書に限定した場合、「2：強い興味」の割合は25.0%となり、推定評価値を限定しない場合の同割合11.1%より高くなることが分かる。だがその代償として推薦する図書の数は198冊から32冊に減って

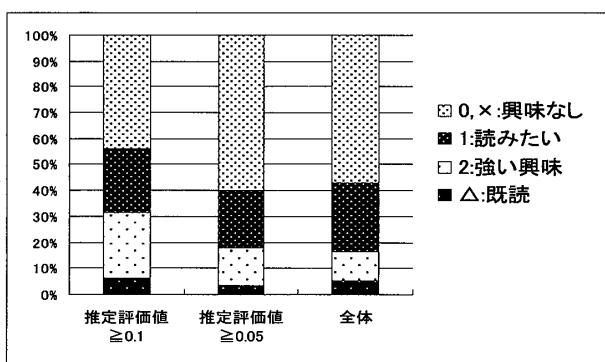


図7：協調フィルタリングにおける推定評価値別の結果

表9：協調フィルタリングにおける推定評価値別の結果

	△：既読	2：強い興味	1：読みたい	0，×：興味なし	合計
推定評価値≥0.1	2(6.3)	8(25.0)	8(25.0)	14(43.8)	32
推定評価値≥0.05	3(3.3)	13(14.4)	20(22.2)	54(60.0)	90
全体	11(5.6)	22(11.1)	52(26.3)	113(57.1)	198

表10：アソシエーションルールにおける確信度・支持度別の結果

	△：既読	2：強い興味	1：読みたい	0，×：興味なし	合計
条件A	1(2.3)	10(22.7)	14(31.8)	19(43.2)	44
条件B	1(1.4)	17(24.3)	26(37.1)	26(37.1)	70
全体	6(4.3)	26(18.6)	52(37.1)	56(40.0)	140

いる。さらに表3の結果と合わせることで、推定評価値が0.1以上の図書に限定しても、肯定的評価の割合はアソシエーションルールのそれに及ばないことが分かる。

6.5 アソシエーションルールにおける確信度・支持度別の結果

アソシエーションルールが推薦した図書の確信度は平均0.34、最大値1、最小値0.04、中央値0.25で、同じく支持度は平均2.82E-06(=0.00000282)、最大値1.13E-05、最小値1.41E-06、中央値2.82E-06であった。そこで推薦する図書を以下の条件A、Bに合うものにそれぞれ限定し、評価結果をまとめた。条件Aは「確信度が0.2以上、支持度が2.82E-06」、条件Bは「確信度が0.1以上、支持度が2.82E-06」である。結果は表10、図8のようになった。全般に条件A、Bのように推薦図書を限定したとしても肯定的評価の割合はあまり変わらないことが分かる。

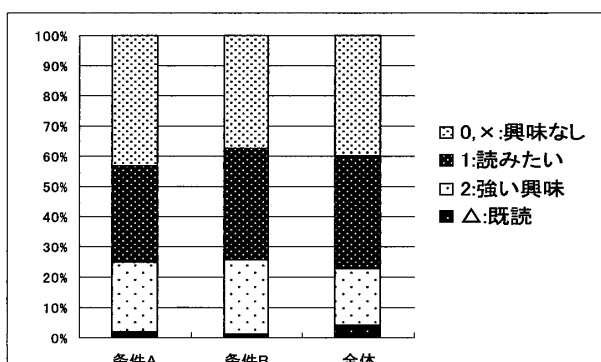


図8：アソシエーションルールにおける確信度・支持度別の結果

September 2012

6.6 分野別の結果

図書推薦の評価結果は主題分野によって異なるかもしれない。そこで図書に付与されている NDC の類を分野とみなし、以下の3種類の集計を行った。即ち、(1)これまで借りた図書において最も多かった分野別、(2)現在借りたい図書の分野別、(3)これまで借りた図書において最も多かった分野と現在借りたい図書の分野が一致しているか否か、の3種類である。以下ではそれぞれについて述べる。

まず被験者ごとに、これまで借りた図書において最も多かった分野を調べたところ、最も多かったのは0類、次は3類であった。それぞれに該当する被験者の推薦評価結果をまとめると表11、表12のようになった。⁴⁴⁾0類3類以外は区別せずにまとめたところ表13のようになった。表11と表13を比較すると、これまで借りた図書において最も多かった分野が0類の者と、0類3類以外の者として評価結果が大きく変わるようには見えない。例えば「2：強い興味」と評価した者の割合は、表11では Amazon、アソシエーションルール、協調フィルタリングそれぞれ 26.7%、19.2%、11.9% となっており、この順に割合が高いが、表13でもそれぞれの割合は 42.1%、26.9%、6.7% となっており、割合の大小は変わらない。表12を見ると「0、×：興味なし」の割合はアソシエーションルールが 55.6% と最も高く、次は協調フィルタリングの 41.7% となっているが、サン

辻ほか：図書館の貸出履歴を用いた図書推薦システムの有効性検証

プルが少ないため統計的に有意なものではない。以上のことから、推薦評価結果が「これまで借りた図書において最も多かった分野」によって大きく異なる可能性は低いように思われる。

次に現在借りたい図書の分野を調べたところ、先ほど同様最も多かったのは0類、次は3類であった。それぞれに該当する被験者の推薦評価結果をまとめると表14、表15のようになった。0類3類以外は区別せずにまとめたところ表16のようになった。これらの表から「借りたい図書の分野」によって推薦評価結果が大きく異なる可能性は低いように思われる。

これまで借りた図書において最も多かった分野と、現在借りたい図書の分野が一致していない場合、その被験者はこれまでと興味関心が変わっており、前者に依存して図書を評価する協調フィルタリングの推薦結果は低く評価されるかもしれない。そこで被験者を前者と後者が一致しているグループと一致していないグループに分け、推薦評価結果をまとめてみた。結果は表17、表18のようになった。表17から一致しているグループでは協調フィルタリングに関する肯定的評価は 47.8% (=7.8+12.2+27.8) であることが分かる。一方、表18から一致していないグループでは同評価は 38.9% (=3.7+10.2+25.0) であることが分かる。両割合の間には有意水準 0.05 では差が認められないものの、一致していないグループの方が肯定的評価の割合は一見低いことが確認さ

表11：これまで借りた図書において最も多かった分野が0類の者の結果

	△：既読	2：強い興味	1：読みたい	0、×：興味なし	合計
Amazon	11(14.7)	20(26.7)	27(36.0)	17(22.7)	75
アソシエーション	5(6.4)	15(19.2)	26(33.3)	32(41.0)	78
協調フィルタリング	9(7.1)	15(11.9)	32(25.4)	70(55.6)	126

表12：これまで借りた図書において最も多かった分野が3類の者の結果

	△：既読	2：強い興味	1：読みたい	0、×：興味なし	合計
Amazon	1(6.7)	2(13.3)	9(60.0)	3(20.0)	15
アソシエーション	0(0.0)	0(0.0)	8(44.4)	10(55.6)	18
協調フィルタリング	2(8.3)	5(20.8)	7(29.2)	10(41.7)	24

表13：これまで借りた図書において最も多かった分野が0類3類以外の者の結果

	△：既読	2：強い興味	1：読みたい	0、×：興味なし	合計
Amazon	2(5.3)	16(42.1)	14(36.8)	6(15.8)	38
アソシエーション	2(3.8)	14(26.9)	22(42.3)	14(26.9)	52
協調フィルタリング	1(1.7)	4(6.7)	20(33.3)	35(58.3)	60

表14：現在借りたい図書の分野が0類の者の結果

	△：既読	2：強い興味	1：読みたい	0, ×：興味なし	合計
Amazon	5(6.6)	23(30.3)	33(43.4)	15(19.7)	76
アソシエーション	1(1.5)	13(19.7)	28(42.4)	24(36.4)	66
協調フィルタリング	7(6.9)	13(12.7)	26(25.5)	56(54.9)	102

表15：現在借りたい図書の分野が3類の者の結果

	△：既読	2：強い興味	1：読みたい	0, ×：興味なし	合計
Amazon	4(50.0)	0(0.0)	3(37.5)	1(12.5)	8
アソシエーション	1(6.3)	1(6.3)	7(43.8)	7(43.8)	16
協調フィルタリング	2(8.3)	1(4.2)	11(45.8)	10(41.7)	24

表16：現在借りたい図書の分野が0類3類以外の者の結果

	△：既読	2：強い興味	1：読みたい	0, ×：興味なし	合計
Amazon	4(10.5)	12(31.6)	13(34.2)	9(23.7)	38
アソシエーション	4(6.9)	12(20.7)	17(29.3)	25(43.1)	58
協調フィルタリング	2(2.8)	8(11.1)	15(20.8)	47(65.3)	72

表17：最も多かった主題分野と現在借りたい図書の分野が一致している者の結果

	△：既読	2：強い興味	1：読みたい	0, ×：興味なし	合計
Amazon	5(8.8)	16(28.1)	22(38.6)	14(24.6)	57
アソシエーション	3(5.6)	7(13.0)	24(44.4)	20(37.0)	54
協調フィルタリング	7(7.8)	11(12.2)	25(27.8)	47(52.2)	90

表18：最も多かった主題分野と現在借りたい図書の分野が一致していない者の結果

	△：既読	2：強い興味	1：読みたい	0, ×：興味なし	合計
Amazon	8(12.3)	19(29.2)	27(41.5)	11(16.9)	65
アソシエーション	3(3.5)	19(22.1)	28(32.6)	36(41.9)	86
協調フィルタリング	4(3.7)	11(10.2)	27(25.0)	66(61.1)	108

れた。学生はあまり関心がない科目(分野)のレポートや期末試験のために図書を借りることも多い。過去に借りた図書と現在の関心が一致しないケースは多くその意味で注意が必要と思われる。

7. 考察

前章で述べたように、推薦された図書に対する被験者の肯定的評価は全般に Amazon, アソシエーションルール, 協調フィルタリングの順に高かった。これは被験者を学年別に見ても、また図書の分野別に見ても言えた。アソシエーションルールと比較した場合、協調フィルタリングには以下の3つの問題点がある。

(1)第1章で述べたように、協調フィルタリングを

行うには利用者一人一人の貸出履歴を保持し続ける必要がある。一方、アソシエーションルールと一緒に貸し出された図書の記録を保持するだけでよく、個々人の貸出履歴を保持し続ける必要がない。従って協調フィルタリングに比べ、アソシエーションルールの方が個人情報漏洩の危険性が少ない。

(2)少なくとも本研究が採用した協調フィルタリングとアソシエーションルールとでは、前者の方がはるかに計算コストがかかる。

(3)貸出冊数が21冊以上のグループと20冊以下のグループとでは、協調フィルタリングは後者において推薦パフォーマンスを低下させた。貸出冊数が少ない利用者に対して協調フィルタリング

September 2012

は有効に機能しない可能性がある。

以上のことから、図書館において貸出履歴を用いた図書推薦システムを構築・提供するならば、推薦パフォーマンスが相対的に低く、かつ(1)(2)(3)のような弱点を抱える協調フィルタリングを用いるのは得策ではない。それよりはアソシエーションルールを用いた方が有効・安全・高速に推薦が行えると言える。さらにそうしたアソシエーションルールより Amazon の方が推薦パフォーマンスが高かったこと、Amazon は貸出履歴を用いる必要が一切ないことを考えると、Amazon を図書館における図書推薦システムに組み込むことも考えられる。Amazon は自身の推薦手法を公開していないが、各図書のページに BOOK データベースの内容紹介文を掲げていることから、推薦に当たっては購入履歴以外にそれら紹介文の類似度も参考にしていることは明らかであろう。コンテンツ(この場合内容紹介文)の類似度に基づく推薦は一般に有効であることから、図書館が Amazon を超える推薦を行おうと思ったら何らかの情報源によって図書の内容も把握すべきと思われる。だが BOOK データベースは非商用利用で毎月1回更新の契約でも年間441,000円かかる⁴⁵⁾。従って図書の内容に関して、BOOK データベース以上に有効な情報源がなく、かつ上記のような金額を図書の推薦にかけられないならば、図書館は Amazon の推薦を超えられない可能性が高い。以上のことから、(a)Amazon の推薦システムはブラックボックスであり今後どのような変更があるかわからないこと、(b)同社が API の提供をやめる可能性が常にあること、(c)絶版になるなどして市場に流通しておらず図書館でしか利用できない図書が推薦されないこと、などは考慮に入れねばならないものの、図書館が図書の推薦において Amazon も併用することは検討に値するように思われる。

8. おわりに

本研究では、(1)T 大学附属図書館の貸出履歴を用いた協調フィルタリング、(2)同データを用いたアソシエーションルール、(3)Amazon、の3つによって被験者に図書を推薦し、有効性を比較検証した。結果、(3)(2)(1)の順にパフォーマンスが高いことを確認した。

今後の課題としては以下の5つが考えられる。即ち、(1)協調フィルタリングに関する各種パラメー

辻ほか：図書館の貸出履歴を用いた図書推薦システムの有効性検証

タを変えた推薦、(2)主題(コンテンツ)に基づく推薦、(3)学年や時期に基づいて貸出履歴を細分した上での推薦、(4)他の大学図書館における貸出履歴の利用、(5)公共図書館における貸出履歴を用いた推薦、の5つである。

まず(1)だが、本研究では原田・増田の手法に基づいて、標本利用者が借りた図書には重み1を付与し、Resnick らの手法との整合性から借りていない図書には重み0を付与した。この点については様々な変更が考えられる。今後は標本利用者ごとに最も多く借りていた NDC の図書に大きな重みを付与するといった変更を行い、その有効性を検証したい。また本研究では原田・増田に基づいて、200冊以上500冊未満の図書を借りた利用者に標本利用者を限定した。こうした限定を行わなかった場合どうなるかも検証したい。⁴⁶⁾

次に(2)の主題に基づく推薦としては、各図書の NDC、タイトル中のキーワード、あるいは BOOK データベースが挙げる内容、といった主題に基づく推薦が考えられる。(3)としては大学生の成長や学期末における利用か否かといった細かい違い・変動を考慮することが考えられる。だが少なくとも本研究が対象とした大学図書館では、利用される図書の異なり数は学生数に比べてかなり大きく、貸出履歴の細分化は安定した結果を生まない可能性が高い。そのため(4)のように他の大学図書館における貸出履歴を利用することが考えられる。もし日本の大学図書館が協力し合い、利用者のプライバシーを守る形で貸出履歴を利用できるようにしたならば、(3)のように細かい条件で貸出履歴を分けても十分なサンプルを確保でき、より適切な推薦が行えるようになるかもしれない。最後に(5)だが、公共図書館では本研究と異なる傾向が見られるかもしれない。利用者のプライバシーを守る形で公共図書館から貸出履歴を得ることが出来たならば、本研究と同様の検証を行ってみたい。

以上の方向で今後研究を進め、図書館において最も有効な図書推薦のあり方を検証していきたい。

謝辞

実験に協力して下さった被験者の皆様、貸出履歴データを提供して下さった T 大学附属図書館の皆様に、この場を借りて御礼申し上げます。本当にありがとうございました。

注・引用文献

- 1) 安東奈穂子ら「電子図書館と利用者のプライバシー：履歴・個人情報の保護と利用の両立を目指して」『デジタル図書館』(30), 2006, p.62-71.
- 2) 岡本真「Web2.0時代の図書館-Blog, RSS, SNS, CGM」『情報の科学と技術』56(11), 2006, p.502-508.
- 3) 渡邊斉志「知的自由の陥穽：利用情報保護思想が公立図書館に及ぼす影響の分析」『Library and Information Science』(58), 2007, p.103-115.
- 4) 佐浦敬之, 辻慶太「公共図書館における利用履歴の活用に関する意識調査」『第57回日本図書館情報学会研究大会発表要綱』, 2009, p.1-4.
- 5) 小野永貴, 常川真央「Web時代にあるべき未来の図書館サービスの胎動：貸出履歴の議論を超えた Shizuku2.0の実現へ」『情報管理』53(4), 2010, p.185-194.
- 6) 成田市立図書館. <<http://www.library.narita.chiba.jp/>>. [引用日：2012-03-09]
- 7) 原田隆史「図書館の貸出履歴を用いた図書の推薦システム」『デジタル図書館』(36), 2009, p.22-31.
- 8) 原田隆史, 増田浩佑「貸出記録を用いた図書推薦システムにおける重み付けの変更」『デジタル図書館』(38), 2010, p.54-66.
- 9) 協調フィルタリングについては、顧客の属性(例えば性別, 年齢, 趣味など)ごとに履歴を記録しておき, それのみを使って推薦を行うタイプもある。このタイプは記録が漏洩しても比較的ダメージが少ないメリットがあるが, 個々人に特化した推薦は行えない。本研究における協調フィルタリングにこのタイプは含めないこととする。
- 10) 原田隆史, 増田浩佑, 前掲8)。
- 11) Amazon. <<http://www.amazon.co.jp/>>. [引用日：2012-03-09]
- 12) 斉藤国博, 中條将典「おすすめ商品を見つける仕組み：「Amazon.co.jp」(アマゾン ジャパン) など：どうなっているの？あのソフトの仕組み」『日経 ITPro』2006. <<http://itpro.nikkeibp.co.jp/article/COLUMN/20061012/250605/?ST=develop&P=3>>. [引用日：2012-03-09]
- 13) 原田隆史, 増田浩佑, 前掲8)。
- 14) Paul Resnick et al., "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, 1994, p.175-186.
- 15) 神瀧敏弘「推薦システムのアルゴリズム(2)」『人工知能学会誌』23(1), 2008, p.89-103.
- 16) 原田隆史, 前掲7)。
- 17) 原田隆史, 増田浩佑, 前掲8)。
- 18) Keita Tsuji et al., "Use of Library Loan Records for Book Recommendation," *Proceedings of the International Conference on Integrated Information (IC-ININFO 2011)*, 2011. (No Pagination)
- 19) 原田隆史, 前掲7)。
- 20) 原田隆史, 増田浩佑, 前掲8)。
- 21) Keita Tsuji et al., op. cit. 18).
- 22) Colleen Whitney and Lisa Schiff "Melvil Recommender Project: Developing Library Recommendation Services," *D-Lib Magazine*, 12(2), 2006. <<http://www.dlib.org/dlib/december06/whitney/12whitney.html>>. [引用日：2012-03-09]
- 23) Saurabh Shirgaonkar et al., "Application of Improved Apriori in University Library," *International Conference and Workshop on Emerging Trends in Technology (ICWET 2010)*, 2010, p.535-540.
- 24) Chia-Chen Chen and An-Pin Chen "Using Data Mining Technology to Provide a Recommendation Service in the Digital Library," *The Electronic Library*, 25(6), 2007, p.711-724.
- 25) Yongcheng Luo et al., "A Privacy-Preserving Book Recommendation Model Based on Multi-Agent," *Proceedings of the 2009 Second International Workshop on Computer Science and Engineering*, 2009, p.323-327.
- 26) 神瀧敏弘「推薦システムのアルゴリズム(1)」『人工知能学会誌』22(6), 2007, p.826-837.
- 27) 神瀧敏弘, 前掲15)。
- 28) 神瀧敏弘「推薦システムのアルゴリズム(3)」『人工知能学会誌』23(2), 2008, p.248-263.
- 29) T 大学は日本有数の研究系総合大学である。本研究で得られた傾向は他のタイプの大学では観察されないかもしれない。この点の検証は今後の課題としたい。
- 30) その他の利用者は専攻分野など各種属性が把握しづらいため除外することにした。
- 31) 推薦に用いる情報の量をなるべく等しくする為, 協調フィルタリングにもこの1冊は「既に借りた1冊」として用いた。
- 32) 原田隆史, 増田浩佑, 前掲8)。
- 33) Paul Resnick et al., op. cit. 14).
- 34) 神瀧敏弘, 前掲15)。
- 35) 原田隆史, 増田浩佑, 前掲8)。
- 36) 福田剛志ら『データマイニング』共立出版, 2001, p.19-32. (データサイエンス・シリーズ3)。
- 37) 金明哲「フリーソフトによるデータ解析・マイニング第40回：アソシエーション分析(1)」『ESTRELA』(152), 2006, p.52-57. <http://mj.in.doshisha.ac.jp/R/200611_40.pdf>. [引用日：2012-03-09]
- 38) これらの値は経験的なものである。いくつかの値を試し, その中で最も有効と思われる値を採用した。用いるデータが変わった場合はそれに合わせて修正すべき値と思われる。
- 39) ただし本研究では21冊以上から成るトランザクションは

September 2012

研究例会案内

ルール抽出に用いなかった。これは何らかの特殊な事情で借りられたと思われ、またルール抽出に用いるとその計算量を著しく増やす為である。なお上記のアソシエーションルールが6冊提示しなかった場合は、提示した5冊以下の図書のみを被験者に推薦した。

40) Amazon が6冊提示しなかった場合は、提示した5冊以下の図書のうち T 大学附属図書館が所蔵している図書のみを推薦した。

41) Amazon Product Advertising API.

〈<https://affiliate.amazon.co.jp/gp/advertising/api/detail/main.html>〉. [引用日: 2012-03-09]

42) 原田隆史, 増田浩佑, 前掲8).

43) 池田央ら『統計ガイドブック』新曜社, 1989, p.86.

44) なお最も多い類が0類と3類で同数という被験者がいたため合計数は表3のそれとは一致しない。

45) BOOK データベース. 〈<http://www.nichigai.co.jp/dcs/index3.html>〉. [引用日: 2012-03-09]

46) ただし現在得ている部分的な結果からは、推薦パフォーマンスは若干向上するもののアソシエーションルールや Amazon には及ばないこと、計算量が膨大になり、かなりの計算資源を必要とすることが分かっている。

◆研究例会案内◆

〈第292回〉

テーマ: 矯正施設における読書環境について: 図書館との関わりを中心に

発表者: 日置将之氏(大阪府立中央図書館)

日 時: 2012年9月29日(土) 14:00~16:00

会 場: 大阪市立弁天町市民学習センター 第3研修室

大阪市港区弁天1-2-2 オーク2番街
7階 Tel: 06-6577-1430

地下鉄中央線「弁天町」西口2A出口より徒歩3分。JR 環状線「弁天町」駅下車北口より徒歩3分

要 旨: 矯正施設とは、法務省が所管する刑務所や少年院等の総称である。これらの施設には図書室等が備え付けられており、読書活動も行われているが、施設の性質上、その実情が図書館関係者の目に触れるメディアで紹介されることは少なかった。そのような中、矯正施設内の読書環境整備や図書館との連携充実を目的とした「矯正と図書館

サービス連絡会」が2010年9月に発足している。この団体の発足から約2年が経過し、関連文献も増えていることから、矯正施設の読書環境等に対する図書館関係者の関心は徐々に高まりつつある。そこで本発表では、図書館関係者の関心をさらに深め、何らかのアクションを起こす契機としてもらうため、矯正施設の読書環境や図書館との連携事例等を詳しく紹介し、今後の展望等について述べる。

〈第293回〉

テーマ: 図書館職の記録研究グループからの中間報告: 図書館職のライフ・ライティング『ゾイア!...』をめぐって

発表者: 田口瑛子氏(京都精華大学名誉教授)

日 時: 2012年10月13日(土) 14:00~16:00

会 場: 大阪市立総合生涯学習センター 第5研修室
大阪市北区梅田1-2-2 大阪駅前第2ビル5階 Tel: 06-6345-5000

要 旨: 著者は現在94歳のロシア生まれのユダヤ人で、70歳代後半の著作である。1940年代から80年代にかけて、断続的に図書館員として働いてきた。その職業生活上の焦点の1つは、ベトナム反戦活動にかかわる連邦大陪審で、図書館員として知的自由をまもるため証言拒否をしたかどで投獄されたことである。アメリカ図書館協会知的自由委員会議長としても活躍した。

翻訳を通して得た、わたしたちの記録研究へのヒントなども述べたい。(『ゾイア!: ゾイア・ホーン回顧録, 知る権利を求めて闘う図書館員』田口瑛子訳, 京都図書館情報学研究会, 2012)

さらに、グループメンバーからの評価などを用意する。事務局から山田伸枝, 伊藤昭治氏の記録への反響や今後の方向についても報告する。

研究例会は、どなたでも自由にご参加いただけます。