

図書館の貸出履歴を用いた図書の推薦システム

The book recommendation system using library loan records

原田隆史

慶應義塾大学文学部

〒108-8345 東京都港区三田 2-15-45

E-mail: ushi@slis.keio.ac.jp

大学図書館の貸出履歴 8808 人分の 960078 件を元に、協調フィルタリングにより図書を推薦するシステムを開発し利用者の評価実験も行った。35 人の被験者が、それぞれ 30 冊の推薦された図書の評価を行った結果、評価可能であった図書のうち 55.4%の推薦図書について有用であるという評価が得られ、うち 15.4%については非常に高い評価が得られた。特に利用した図書の数が多い利用者ほど高い評価をする傾向が見られた。また、図書と利用者との対応関係を秘匿するために利用者をグループ化して推薦した場合でも、推薦結果の評価は大きくは下がらなかった。分析の結果、日本の大学図書館での利用によく見られるレポート執筆用の貸出が多いことが評価を下げる大きな要因のひとつであることが明らかとなった。

We have developed the book recommendation system using collaborative filtering, based on 960078 loan records of 8808 users of the academic library. The evaluation experiment was then conducted. In this experiment, each of 35 test subjects evaluated recommended books. The result showed that (1)55.4% of the recommended books were useful, especially 15.4% of them were rated quite high, (2)the more books users borrow, the higher they rated, (3)in case of grouping users in order to conceal the relation between books and users, the result of evaluation did not decrease considerably. The results of analysis indicate that one of the main reasons of decreasing the evaluation was that many loan order were made for assignments, which was often the case in academic libraries in Japan.

1. 図書館における貸出記録の保存

図書館における「貸出記録」とは、図書館において、利用者に貸出された資料が返却されてくるまでの期間、①何が、②誰に、③いつまでの期限で、貸出されているかを記載し、図書館に保管される記録のことである¹⁾。貸出記録は「貸出履歴」と表現されることもあるが、これまで、特に両者の明確な定義は行われていない。貸出記録は「現時点で貸出中の資料の貸出記録」と「資料返却後保存している貸出記録」とは役割も管理上の取り扱いも異なり、分けて考えられることが多い。貸出を行った資料が返却期限を過ぎても返却されない場合、図書館は利用者に対して督促状を郵送する、電話や FAX

などによって返却期限が過ぎている旨を伝えるといったことをする必要があり²⁾、図書館の資料管理上から前者を図書館で記録しておくことは当然とみなされる。それに対して、後者の保存に関しては特に公共図書館では、個人情報保護の観点などから保存すべきでないという意見も根強い³⁾。本研究で述べる「貸出記録」は後者の「資料返却後も保存している貸出記録」を指すものとする。

資料が返却された後も、大学図書館では貸出記録を保存し、利用者の申請に応じて提示を行ったり利用者本人が Web 上で自身の貸出記録を確認したりできるところがある⁴⁾。山口真也が私立大学図書館協会西地区部会に加盟する

大学図書館を対象に行った資料返却後の貸出記録保有状況の調査によれば、調査に回答した198館の87%の大学図書館が資料返却後も貸出記録を残しているとされる⁵⁾。

また、学校図書館においても読書指導の一環として児童生徒の貸出記録を卒業時まで保存し、卒業の際に記念として贈呈するという活動を行っているところも多い⁶⁾。

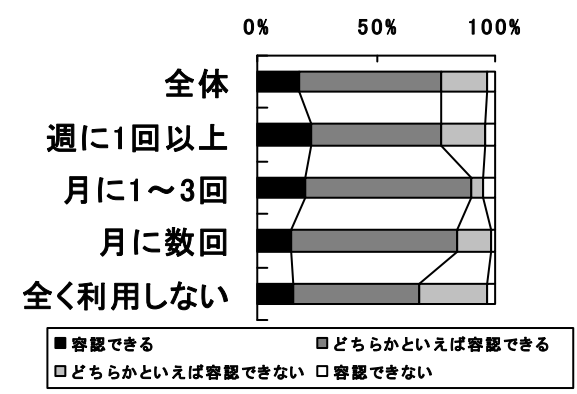
しかし、貸出記録を保存する図書館においても、その利用は前述のような利用者への開示目的および図書館管理などで使用する統計データへの加工という目的以外にはほとんど使われてこなかった。しかし近年、貸出記録をもとに、図書の推薦システムのような積極的な利用者サービスを展開すべきであるという意見も見られるようになってきている⁷⁾。

このようなシステムの構築自体は、技術的にはそれほど大きな問題はない。貸出中の資料であっても返却済みの資料であっても、資料管理やOPAC、館内・外部のデータベースへのアクセス等を担当する統合図書館システム上のレコードであり⁸⁾、単に返却時にデータを削除するかどうかという操作の違いだけに過ぎないともいえる。

問題は、貸出記録の利用に際して利用者の個人的な思想・信条などが漏洩する危険性への配慮であり、利用者自身も利用に慎重であるのではないかと予想される。この点に関しては、図書館の提供するサービスを利用する主体である図書館利用者の、個人情報に対する意識を調査した筑波大学大学院の佐浦敬之による研究がある⁹⁾。

佐浦は図書館の利用者400人を対象に、図書館の利用頻度別の図書館の貸出記録に関する意識調査を行っている。その結果、利用者の75.1%が過去に借りた本の履歴を確認できるサービスを使いたいと回答し、利用頻度を問わず、図書館がそのような利用者サービスを目的に貸出記録を保存することについても7割以上が「容認できる」「どちらかといえば容認できる」と考えていることを明らかにしている(結果を第1図に示す)。

もちろん、貸出記録の保存を利用者が無条件に許容するわけではない。同時に行われた貸出記録の保存をする上での図書館に求める条件についての調査では、「貸出記録から個人が特定されないこと」、「サービスを利用するか否かという選択が利用者に委ねられていること」などが特に強く求められている。いずれにしても、「いかなる場合も貸出記録の保存を認めない」という者は全体の2%とごくわずかであった。



第1図 サービス提供のために貸出記録を保存することの是非(図書館利用頻度別)

2 貸出記録を用いた利用者への図書の推薦

2005年の秋頃から急激に日本でも認知され始めたWeb2.0¹⁰⁾と呼ばれるWebサービスでは、利用履歴を用いた推薦システムが数多く実用化されてきている。

たとえば、Amazonが行う「おすすめ商品」サービスと名付けられた推薦システムは、利用者の購入履歴、商品の評価、持っている商品などのデータをもとに、他の利用者の購入履歴などを用いて興味があると思われる商品を推薦するものである(第2図)¹¹⁾。



第2図 Amazonの「おすすめ商品」サービス

Amazon は推薦手法については明らかにしていないが、このような推薦には協調フィルタリングの手法がよく用いられており、Amazon でも協調フィルタリングの手法を取り入れていることが考えられる¹²⁾。

図書館の資料を対象として、商品(図書)の推薦を行うことへの需要については、佐浦が前述の貸出記録利用の容認とともに調査を行っている⁹⁾。その結果、過半数の利用者が「とても使いたい」「どちらかといえば使いたい」という意見をもっており、図書館でも、図書の推薦サービスを行うことが有効であることが予想される(結果を第1表に示す)。

第1表 貸出記録をもとにした図書推薦システムに関する調査(図書館利用頻度別)

	全体	とても使いたい	どちらかといえば使いたい	どちらかといえば使いたくない	全く使いたくない
全体	400 100.0	47 11.8	164 41.0	128 32.0	61 15.3
週に1回以上	100 100.0	17 17.0	40 40.0	28 28.0	15 15.0
月に1~3回	80 100.0	14 17.5	35 43.8	25 31.3	6 7.5
年に数回	70 100.0	10 14.3	29 41.4	24 34.3	7 10.0
全く利用しない	150 100.0	6 4.0	60 40.0	51 34.0	33 22.0

図書館が貸出記録を利用して同じような図書の推薦を行うことが可能かという点については、岡本真がAmazonの「おすすめ商品」サービスを引き合いに出し、図書館の貸出記録を利用する必要性を強く主張するとともに、その有効性を検討している⁷⁾。岡本は2005年に取次ルートで販売された本が約7億4000万冊であるのに対し、同年の公共図書館の合計貸出点数が約6億1000万冊とほぼ大差ないことをあげ、「貸出と販売を同義にとらえることはできません。しかし、ポイントはいつ誰が何を手にしたかという意味では同じ記録ということです。」と述べている。貸出記録を用いた図書推薦システムを採用するかどうかの是非は別として、その効果や実装方法などの検討を行う価値は十分にあり、今後の推薦システムに関する研究の必要性は高いといえよう。

しかし、現在まで貸出記録を用いた図書の推薦などのサービスに関する有効性の検討も、実現に向けた研究もほとんど行われていない¹³⁾。

特に実用システムの開発を視野にいれて、システムの構築を行っている例は皆無に近い。わずかに、貸出記録の持ち得るポテンシャルを引き出すとする試みの一環として、次世代図書館情報システムを開発しようとするProject Shizukuの活動が見られる程度である。Shizukuは、IPA未踏ソフトウェア創造事業(平成19年度第I期末踏ユース)の支援を受けて筑波大学図書館情報専門学群の学生が開発を行っている図書館システムで、「貸出記録のビジュアライズ(視覚化)」や「貸出記録に基づく類似ユーザのサジェスト」の機能を実現しようとしている¹⁴⁾¹⁵⁾。

本研究は、実際の図書館の貸出記録をもとに、大量のデータを用いた図書の推薦を行い、その有効性と問題点について検討を行うことを目的としている。

3. 貸出記録を用いた図書の推薦実験

3.1 協調フィルタリング

本実験では、ある私立大学図書館の実際の貸出記録約90万件を用いて、協調フィルタリングの手法を用いて図書の推薦を行う。

協調フィルタリングは、大量の情報の中から有用な情報を選び出す情報フィルタリングの一手法である¹⁶⁾。ユーザの嗜好を過去の行動という形で記録し、そのユーザと似たような行動を取っているユーザの嗜好情報をもとに、ユーザの嗜好を推測する。協調フィルタリングの特徴は、未知アイテムの有用度推定において、他ユーザの意見(既知アイテムに対する有用度評価値)を利用する点にある¹⁶⁾。

協調フィルタリングにおいて利用者の嗜好データを獲得するアプローチには明示的と暗黙的の二種類があり、利用者に意図的に商品などを評価させるかたちをとるようなものが明示的、利用者の行動などから間接的に嗜好を読み取るものが暗黙的アプローチであるとされる¹⁷⁾。本実験では、図書館の利用記録をもとにして推薦を行う暗黙的アプローチを採用した。利用記録をもとにしているものであり、図書館利用者に図書の有用性を評価してもらう

ことは不可能である。そこで、どの図書についても貸出があれば重み1として推薦を行った。

図書の推薦に用いたシステムとしては、2005年に Stéphane Droux により開発・公開された協調フィルタリング手法に基づく情報推薦システムである Vogoo PHP Pro v2.2¹⁸⁾ をもとにして、利用データにあわせたカスタマイズや処理速度の向上などのための修正を行って使用した¹⁹⁾。

3.2 貸出記録を用いた図書推薦実験

本研究で貸出記録を推薦データの決定に用いた大学図書館では、奉仕対象となる学生数が約3000名、学生の専攻する分野は人文科学、社会科学、自然科学と多岐に渡っている。蔵書規模は、和洋図書あわせて約65万冊、年間入館者数は約30万人である。本実験では1991年から2007年までの期間の貸出記録からリザーブブック、雑誌、また延長手続きによる貸出し、長期貸出、図書以外の貸出などを除外した学部生及び大学院生の貸出記録8808人分、960078件を利用した。貸出された異なり図書数は200405冊(蔵書の約31%)で、貸出図書1冊あたりの貸出回数は4.79回となる。

また、貸出記録をもとにした図書の推薦の判定を行う被験者としては、利用者本人に貸出記録の提供を行うサービスがある大学に在籍する東京都内の大学生で、公開できる貸出図書数が21冊以上ある計35名とした。被験者の専攻する分野は、人文科学(文学、史学、哲学など)10名、社会科学(経済学、政治学、法学、教育学、商学など)10名、自然科学(理学、工学など)10名、および複合領域として図書館情報学5名となるようにした。被験者自身が貸出記録を入手し、その中から公開することを許諾した図書のリストをもとにして図書の推薦を行った。

図書推薦のもととなる貸出記録を提供いただいた大学とは別の大学に所属する被験者も含まれていた。その被験者については、貸出記録中の図書とISBNが一致する図書のみを対象とした。

被験者の分野別貸出冊数別の人数を、分野別の平均貸出冊数を男女比とともに第2表に示

す。第2表に示すように、平均貸出冊数は、各分野で大きな違いはなかった。

第2表 被験者の属性

	合計人数	貸出 21~30冊	貸出 31~40冊	貸出 41~60冊	貸出 61冊以上	平均貸出 冊数	男女比
人文科学	10	1	5	2	2	48.3	M4:F6
社会科学	10	3	2	4	1	43.3	M5:F5
自然科学	10	3	2	2	3	52.6	M9:F1
図書館情報学	5	1	1	1	2	50.6	M3:F2
合計	35	8	10	9	8	48.4	M21:F14

各被験者に、推薦結果の図書上位30冊に対して、興味の度合いを「強い興味がある」「読んでみたい」「読みたいとは思わない」という3件法で評価してもらった。利用者に評価を求める際には、まず「書名」「著者名」「出版者名」「出版年」を提示し、それだけでは内容の判断がつかないものについては、可能な範囲でオンライン書店へのリンクを設定して、内容や書評などが参照できるようにした。それでも内容が把握できないなどで評価ができないものは「評価不能」とした。

また、一部の被験者については、どのような基準で評価したかについて、評価実験後にインタビューを行った。

3.3 推薦された図書の評価

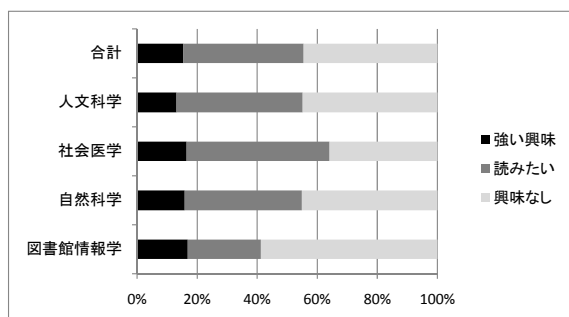
推薦された図書に対する被験者の評価結果として、興味の度合いに属する図書の平均冊数を、被験者の専攻別に第3表に示す。第3表に示すように、推薦された30冊の図書のうち、平均4.8冊については内容の判断ができないなどの理由で評価不能と判定された。なお、第3表で「強い興味がある」「読んでみたい」「読みたいとは思わない」の欄に平均冊数とともに示した百分率は被験者が評価可能であると判断した図書中での割合を意味している。

第3表に見られるように、被験者の平均で、評価判定ができた25.2冊のうち15.0冊の図書が「強い興味がある」または「読んでみたい」と判定された。これは、評価不能である図書を除いた場合55.4%にあたる。被験者の分野別の

評価の値の割合を第3図に示す。

第3表 被験者分野別推薦結果の評価

	強い興味	読みたい	興味なし	評価不能
合計	3.9(15.4%)	10.1(40.0%)	11.2(44.6%)	4.8
人文科学	3.3(13.1%)	10.6(42.1%)	11.3(44.8%)	4.8
社会医学	4.1(16.5%)	11.8(47.6%)	8.9(35.9%)	5.2
自然科学	3.9(15.9%)	9.6(39.0%)	11.1(45.1%)	5.4
図書館情報学	4.6(16.9%)	6.6(24.3%)	16.0(58.8%)	2.8



第3図 分野別推薦結果の評価
(評価可能図書についての比率)

第3表および第3図に示すように、「強い興味がある」と判定された推薦図書の比率は、各分野で13～17%と大きな差は見られなかったが、「読んでみたい」と評価された図書の比率は分野によって大きく異なっていた。その結果、社会科学分野では約2/3の推薦図書について「読んでみたい」以上の評価が得られたが、図書館情報学分野については約40%にとどまっていた。人文科学および社会科学分野については、50%強という評価であった。ただし、被験者の数が少ないこともあり、カイ二乗値検定の結果からは有意差は見られないという結果であった。

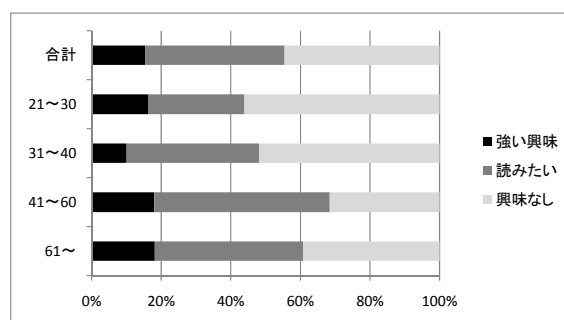
これらの結果は、主として推薦もとのデータとして使用した大学図書館の奉仕対象とする学生および蔵書構成に大きな影響を受けていると考えられる。すなわち、社会科学分野の学生および蔵書数が多く、自然科学分野の学生および蔵書は比較的少ないことが影響している可能性がある。特に、図書館情報学分野の被験

者の評価が低い原因としては、推薦のための元データとして使用した大学に、図書館情報学を専攻する学部・学科が存在しておらず、貸出記録中に学習や研究に適切な図書が少なかったことが大きな原因であると考えられる。

次に、被験者の利用した図書の貸出冊数ごとの推薦された図書に対する被験者の評価結果を第4表に示す。第4表でも百分率は被験者が評価可能であると判断した図書中での割合を示している。また、貸出可能図書についての評価の割合を第4図に示す。

第4表 被験者分野別推薦結果の評価

	強い興味	読みたい	興味なし	評価不能
合計	3.9(15.4%)	10.1(40.0%)	11.2(44.6%)	4.8
21～30冊	4.0(16.3%)	6.8(27.6%)	13.8(56.1%)	5.5
31～40冊	2.5(10.0%)	9.5(50.4%)	12.9(51.8%)	5.1
41～60冊	4.6(18.0%)	12.8(50.4%)	8.0(31.6%)	4.7
61冊以上	4.8(18.2%)	11.1(42.6%)	10.3(39.2%)	3.8



第4図 貸出冊数別推薦結果の評価
(評価可能図書についての比率)

第4表および第4図に示すように、貸出冊数が多い被験者ほど推薦された図書に対する評価が高いという結果が得られた。特に40冊以下の貸出冊数である被験者と41冊以上の貸出冊数である被験者の評価の差は大きかった。第4表および第4図で、61冊以上の貸出冊数の場合が41～60冊の貸出の被験者と比較して評価が低くなっているが、これは図書館情報学を専攻する被験者で貸出冊数の多かった2名の評価が非常に低かったことの影響が大きく、この

2名を除いた場合には41~60冊の場合とほぼ同じ評価であった。

3.4 推薦図書に関するインタビュー調査

推薦された図書についてのインタビューを行った結果、「意識していなかった（知らなかった）が、読みたいと思う図書があった」「研究に役立つような図書があった」などの肯定的意見がある一方、ほしい図書がほとんど推薦されなかったという被験者も存在した。

多くの被験者が推薦される図書には、過去に借りた図書と類似の主題または関連する図書だけではなく、何らの関連性もない図書が多く含まれていたことを指摘している。たとえば、被験者A(西洋史学専攻男性)では推薦された30冊中の19冊、被験者B(法学専攻女性)では30冊中の18冊が過去に借りた図書と関連する図書であり、残りが関連しない図書であった(過去に借りた図書と関連するかどうかは被験者の判定による)。

被験者のうち7名(人文科学, 社会科学, 自然科学各2名, 図書館情報学1名)に推薦された図書210冊中135冊が過去に借りた図書と関連する図書, 75冊が関連しない図書であった。

これら過去に借りた図書と関連する図書が表示される数と、推薦図書の評価との関係はほとんど見られなかった。たとえば、被験者Aと被験者Bとは関連する図書が推薦された数はほぼ同じであるが、被験者Aについては19冊中「強い興味がある」が6冊、「読んでみたい」が1冊だけであった。しかし被験者Bについては18冊中13冊が「強い興味がある」と評価された。このような評価の違いは他の被験者にも多く見られた。

この原因について、実験後のインタビュー調査で以下の3点を指摘する意見が多かった。

まず、利用記録中の図書に関連しているが興味がない図書として、レポート作成のための一時的な利用のために借りた図書と関連するものがあげられる。たとえば、利用者Aは、過去にアラブ地域やイスラムに関する図書を複数冊借りはしたが、その分野は利用者Aの継続的な研究対象ではなく、個人的な興味もないもの

で、レポート作成のための一時的なものだったと述べている。

もうひとつの理由としては、興味の対象が変わってしまったことがあげられる。たとえば、大学入学当初に借りた図書と関係するものについて、既に自分の関心が変わってしまっていたということを指摘する人もいた。先の被験者Bへのインタビューでは、「現在の自分の関心に近い図書の推薦が有効であった」という意見があげられており、推薦された図書が最近(半年以内)利用した図書との関連の強さが高い評価に結びついていることを示している。

さらに、個人の属性よりも個人がおかれた状況などに応じて必要性がきまる図書の存在があげられる。たとえば「地球の歩き方」のような旅行ガイドブックは、その地域への旅行予定が無い人間に推薦されてもあまり意味がない類の図書であると考えることができる。個人の属性に応じた推薦には意味があっても、状況に関わる図書については推薦対象からはずすなどの検討が必要となろう。

3.5 推薦サービスそのものに対する意見

インタビューで図書の推薦サービスへの評価についても求めたところ、概ね好意的な意見が多く聞かれた。代表的な好意的意見は以下のものであった。

- ・関係のありそうな図書を推薦してもらえると探す手間が省けるので、非常に助かる
- ・自分の力だけでは探せない、見つけられないような図書が推薦されるので、便利だ
- ・他の利用者がどんな図書を借りているのかわかるのが面白い

この中で、特に多く得られたのが「面白い」という意見であった。推薦された図書を借りるか借りないか別として、他の利用者がどのような図書を利用しているのかが知れるのは、面白いのだと言う。図書の推薦は、実用だけでなく、娯楽性も兼ね備えているのかもしれない。

否定的な意見は少なかったものの、以下のものが挙げられた。

- ・必要な図書は自分で探すことができるので、このようなサービスは邪魔である

- ・貸出記録が保存されることを快く思わない
- ・あれば便利だと思うが、なくてもさほど困らない

また、このようなサービスへの要望として「推薦する際には、図書を推薦した理由を明らかにしてほしい」「強制的なサービスであって欲しくない」などの要望もあった。

4. 推薦元データを加工した推薦

前述のように、公共図書館を中心に貸出記録を保存するかどうかについては議論がある。その大きな要因としては、貸出記録からの個人情報流出の可能性と、貸出記録が流出したことによる思想信条の漏洩の問題が指摘される。

一方、人間が図書館を利用する際には1つの主題に関する図書だけを利用するわけではなく、多くの主題の図書にまたがって利用することになる。このような状況を考えた場合、データ量が十分に多ければ、複数人をまとめて1つのグループとして推薦を行った場合でも、推薦結果の有用性を確保できる可能性も予想される。そこで本研究では、貸出記録の匿名性を高めるために後述するグループ化の手法を用い、グループ化しない場合の推薦結果と比較することとした。グループ単位での図書の利用データのみを図書館に保存し、これを推薦に利用することで情報漏洩に対する対策となりえるかどうかを検討するためである。

4.1 実験手順

まず、大学図書館の貸出記録中に含まれる、利用者番号(学生 ID)を用いて MD5 ハッシュ値を算出した。利用者番号は利用者を一意に定める要素であり、8808 人程度では生成されたハッシュに重複はないと考えられる。次に 128 ビットのハッシュ値から上位 117 ビットを削除し、各利用者を残りの 11 ビットのみで表現することにより、利用者を 2048 通りのグループに分けた。この1つのグループを1人の利用者として扱って、前章と同様の手法で図書の推薦と結果の評価を行った。

グループ化することで約 4 名の貸出記録が混じったデータを元にした推薦となるため、個

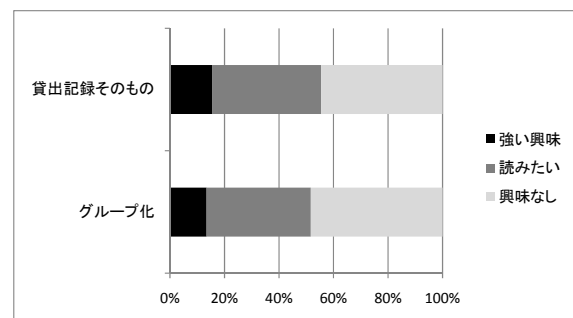
人の利用した図書の履歴を特定することができないようにした状況での推薦が実現できる結果となる。

4.2 グループ化した推薦結果の判定実験

グループ化した貸出記録をもとにした推薦の結果を第 5 表および第 5 図に示す。グループ化しなかった場合と同様に、推薦結果の図書に対する評価は 7 名(人文科学, 社会科学, 自然科学各 2 名, 図書館情報学 1 名)が三件法で行った。

第 5 表 グループ化後の推薦結果の評価

	強い興味	読みたい	興味なし	評価不能
貸出記録そのもの	3.9(15.4%)	10.1(40.0%)	11.2(44.6%)	4.8
グループ化加工後	3.4(13.5%)	9.7(38.2%)	12.3(48.3%)	4.6



第 5 図 グループ化後の推薦結果の評価

第 5 表および第 5 図に示すように、グループ化することで若干の評価の低下はあったものの、貸出記録そのものをもとにする場合と、それほど評価は下がらない結果となった。また、7 名中 2 名の被験者の評価得点は少し上昇した。このことは、グループ化による悪影響がそれほど大きくないことを示しているといえよう。

グループ化をした場合に推薦される図書の内容は、7 名に対して推薦された各 30 冊合計 210 冊中の 34 冊 (16.2%) だけが、貸出記録をもとにして推薦した場合と同じであり、推薦される図書がほとんど入り替わる結果となっていた。ただし、推薦順位 30 位までに含まれてい

た図書が一致した数は少なかったが、貸出記録をもとに推薦した場合の上位 150 位までに含まれていた図書の割合は 210 冊中で 146 冊 (69.5%) に達しており、全く異なる図書が推薦されるのではなく、推薦の順位が大きく変動するという状況であった。

なお、オリジナルの貸出記録をもとにして推薦された図書と、グループ化した図書の両方に含まれる図書が必ずしも高い評価になるわけではなかった。

5. より望ましい図書推薦システムを目指して

本研究では、大量の図書館の貸出記録を用いて図書の推薦を行った場合、評価可能であった図書のうち 55.4% の推薦図書について「読んでみたい」という評価が得られ、うち 15.4% については「非常に興味がある」という高い評価を得ることができた。

特に、推薦を受ける利用者がそれまでに利用していた図書数が多い場合、また推薦の元データとなる図書館での貸出が多い図書を利用していた場合ほど、より適切な図書を推薦することができる可能性を高めることも明らかとなった。その意味で、より適切な図書を推薦するためには、多くの図書館の貸出記録を集めて利用することが有効であろう。

そのためには、推薦を受ける利用者の利用図書と、推薦に用いる貸出記録との照合も重要と考えられる。本研究では特に推薦データのもととなった大学図書館と違う大学に所属する学生については、一部の図書が推薦を利用することができなかった。これは、図書の照合に際して機械化の速度優先のために ISBN による照合のみを行ったためである。今後、実用システムをめざすような場合にはこの作業をどの程度自動化できるかも焦点となろう。

また、本研究のような図書の推薦システムに対して好意的な評価や高い評価をくださった被験者には、図書館の貸出は多いが、書架でのブラウジングをあまり行わず、OPAC で図書を検索して貸出を行う利用パターンをとる人も多くみられた。このような場合、図書を探す時に

は OPAC のみを使って検索し、OPAC で適当な図書が見つからなかった場合は諦めることが多い。そのため、本研究で提示したような図書の推薦システムが興味深く感じられたようであった。

一方、「(図書の推薦サービスが) なくても困らない」と答えた利用者には、頻繁に図書館を利用しており、ブラウジングも日常的に行うという利用者が多く含まれていた。

利用者教育も図書館にとって重要であるが、OPAC から望んだ結果が得られなかったことによって「図書館は使えない」と一方的に判断を下され、図書館を利用するという選択肢が失われることは、図書館にとっても利用者にとっても不幸なことである。図書の推薦サービスは、従来の図書館教育とは別のアプローチから図書館利用のきっかけとなる方法を提示する可能性も持っているといえよう。

ただし、本実験ではいくつかの課題も明らかとなった。特に、推薦に利用する図書データの程度選択するかについては大きな課題が残っている。たとえば、レポートの利用については、一度に複数冊借り、その後は借りられていない分野や主題の図書は、推薦の対象にしないなどの方策が有効かもしれない。レポートでの利用の場合に、どのような図書の利用がなされるかなどについて、さらなる研究が必要となろう。また、大学図書館における図書の推薦においては、貸出時期が古い貸出履歴よりも、新しい貸出履歴の情報を優先的に利用する、旅行ガイドのような図書についての推薦順位を下げるなど、図書の属性に応じた重み付けの検討なども必要となろう。

さらに、グループ化による推薦を行うための貸出記録として利用者情報をハッシュ化し、さらにその一部だけを使用した場合でも、通常の場合とほぼ同程度の水準の推薦を行うことが可能であることは、特に公共図書館などでの実用化を視野に入れる場合、非常に興味深い結果といえよう。

情報の漏洩については単にデータをどのような形で記録するかということだけにとどま

らず、情報システム全体の管理をどのように行うかが重要である。その意味では、データそのものを加工する必要はないという考え方も当然ありえるだろう。しかし現実には、特に公共図書館を中心に、保管体制の不備のために貸出記録を破棄する例が多いことを考えれば、ある程度の評価の低下があったとしても実現性が高くなるという意味でグループ化を行える可能性を示せたことは有意義であると思われる。

近年、Amazon や Google といった個性的で優れたサービスを提供する企業の台頭により、利用者も新しいサービスの展開を期待するようになってきている。このような状況を受けて、図書館も従来のサービスの枠にとらわれない新しいサービスを考えていく必要がある時代になっているといえよう。

本研究のような図書の推薦システムが必要かどうかについての結論は現時点では当然くだすことはできない。しかし、少なくとも有効性についての議論は行ないはじめる必要があるだろう。本研究が、今後の新しいサービスの基礎となり、導入することが有効かどうかという議論へとつながっていけば幸いである。

6. 謝辞

本研究の実験に際して貸出記録を提供くださった大学図書館の関係者諸氏、および被験者として貸出記録を提供いただいた方々に深く感謝いたします。また、実験についてはデータの処理および分析の過程で永友敦子氏、栗原渉氏(慶應義塾大学)の多大なる協力を得ました。ここに感謝いたします。

7. 注・引用文献

- 1) 日本図書館協会図書館ハンドブック編集委員会編. 図書館ハンドブック. 第4版. 東京, 日本図書館協会, 1977, p. 328-330.
- 2) 前島重方, 天野哲雄, 室伏武, 金中利和. 図書館活動. 改訂版. 東京, 樹村房, 1995, p. 89.
- 3) たとえば, 日本図書館協会は「貸出業務へのコンピュータ導入に伴う個人情報の保

護に関する基準(1984年5月25日社団法人日本図書館協会総会議決)」を定めている(<http://www.jla.or.jp/privacy/kasidasi.html>)。そこでは「コンピュータによる貸出しに関する記録は、図書館における資料管理の一環であって、利用者の管理のためではないことを確認し、そのことに必要な範囲の記録しか図書館には残さないことを明らかにして、利用者の理解を得るよう努めなければならない。」とし、6点の基準を示している。その基準の中のひとつとして、「貸出記録は、資料が返却されたらできるだけすみやかに消去しなければならない。」があげられている。

- 4) たとえば, 慶應義塾大学図書館や創価大学中央図書館では, 図書館の利用に関する Web ページ上の説明で図書館の貸出記録を利用者が申請すれば見ることができることが示されている。慶應義塾図書館. 慶應義塾図書館 <FAQ- 回答集>. [2008-12-12], http://www.mita.lib.keio.ac.jp/contact_us/faq/answer/#kashidasi や, 創価大学中央図書館. My season. [2008-12-12], <http://office.soka.ac.jp/library/ms/setumei.htm> など参照。
- 5) 山口真也. 私立大学図書館における個人情報・プライバシー保護 : 貸出記録の管理方法と消去の必要性を中心に. 沖縄国際大学日本語日本文学研究. 2005, Vol. 10, No. 1, p. 1-24.
- 6) 山口真也. 沖縄県学校図書館におけるプライバシー保護の現状に関する調査 : 貸出記録の目的外使用問題を中心として. 沖縄国際大学日本語日本文学研究. 2004, Vol. 9, No. 1, p. 13-68.
- 7) 岡本真. 図書館と Web2.0: Web2.0 時代の図書館-Blog, RSS, SNS, CGM. 情報の科学と技術. 2006, Vol. 56, No. 11, p. 502-508.
- 8) 根本彰. 公共図書館のニューウェイブ: 公共図書館の電子化と公共性. 情報の科学と技術. 2001, vol. 51, no. 7, p. 375-380.

- 9) 佐浦敬之. 利用者からみた利用履歴の活用—意識調査の概要と結果—. 第10回図書館総合展フォーラム「貸出し履歴を利用した新しい利用者支援の展開」2008年11月27日での発表に用いられた資料より
- 10) Tim O'Reilly. O'Reilly Network: What Is Web2.0. 2005-9-30 [2008-11-27], <<http://www.oreillynet.com/lpt/a/6228>>
- 11) Amazon.co.jp. 「おすすめ商品」についてのヘルプ. [2008-11-27]. <<http://www.amazon.co.jp/gp/help/customer/display.html?nodeId=779360>>
- 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 2006-10-26. [2008-11-27]. <<http://itpro.nikkeibp.co.jp/article/COLUMN/20061012/250605/?ST=develop&P=3>>
- 13) 久保山健. 次世代 OPAC を巡る動向: その機能と日本での展開. 2008, Vol. 58, No. 12. p. 602-609.
- 14) 常川真央, 小野永貴, 安西慧, 矢ヶ部光. 利用者のつながりを創り出すコミュニティ指向型図書館システム. 情報処理学会研究報告 - DD(デジタル・ドキュメント). 2008, Vol. 2008, No. 34, p. 1-6.
- 15) Project Shizuku. Project Shizuku ~次世代図書館情報システム. [2008-11-29], <<http://www.shizuku.ac/>>
- 16) 森田昌宏, 速水治夫. 情報フィルタリングシステム. 情報処理. 1996. Vol. 37, No. 8, p. 751-758.
- 17) 神寫敏弘. 情報のフィルタリング: 推薦システム—情報過多時代をのりきる. 情報の科学と技術. 2006, Vol. 56, No. 10, p. 452-457.
- 18) Stéphane Droux. Vogoo - Recommendation Engine & Collaborative Filtering. [2008-12-18], <<http://www.vogoo-api.com/index.php#vogoo>>
- 19) 本実験で使用したデータを用いた図書館の推

薦システムを以下の URL で公開している。試していただければ幸いです。<http://ushi-semi.slis.keio.ac.jp/~ushi/vogoo-pro/recommend.html>