

情報工学実験におけるソーシャルボタン機能の実践とその結果分析

徐海燕[†] 山口亨[‡] 史一華[‡]

[†] 福岡工業大学工学研究科情報工学専攻 〒811-0295 福岡市東区和白東 3-30-1

[‡] 西南学院大学商学部経営学科 〒814-0002 福岡市早良区西新 6-2-92

E-mail: [†] xu@cs.fit.ac.jp, mfm12013@bene.fit.ac.jp [‡] shi@sennan-gu.ac.jp

あらまし データベースと Web プログラミングの技術を学生に実践させるため、徐研究室では、Web モール実験システムを開発し、2003 年度から「情報工学実験Ⅲ、Ⅳ」という 3 年次の科目に利用してきた。システムの機能強化と世の中の流れに沿って利用者同士が店舗を評価し合うソーシャルボタン機能を新規に追加するために、一昨年度から、一から作り直し、今年度の実験から本格的に利用している。例年がない積極的な反応が得られている。本論文では、実験を活性化させた要因を突き止めるために、ソーシャルデータに関する実験結果の解析を行う。利用者への「被アクセス数」と「いいね!数」という異なる因子に属するデータの提供が、重要な一因という結果が得られている。さらに、より教育効果を向上させるための改良点や操作履歴による推薦機能についても報告する。

キーワード ソーシャルボタン機能, 情報工学実験, ランキング, 情報推薦

1. はじめに

実験を通して情報処理の専門知識と応用を身に付けるために、福岡工業大学情報工学科では、3年次の必修科目に「情報工学実験Ⅲ、Ⅳ」という科目が設けられている。学生は複数のグループに分けられ、2週間単位で各教員の教室に回ってそれぞれの実験課題を完成していくというカリキュラムである。著者の徐ではデータベース技術を活用したWebアプリケーション開発という実験課題を担当している。徐研究室では、店情報・商品データ・購買履歴の管理にはデータベース関連知識の活用と実践をするために、HTMLなどのWeb関連知識、PHPとJavaScript言語というプログラミングの関連知識を盛り込むWebショップの作成からWebモールの運用までの開発過程を体験する簡易型の「Webモール実験システム」を構築した。2003年度から実験に利用され始め、好評が得られている[7]。その後、教育効果の向上を図るためには、マニュアルやシステムの改良を重ねてきている。

システムのセキュリティ機能やGUI機能を強化するために、一昨年度からシステムの操作環境を一から作り直し、昨年度から実験に利用され[11]。さらに、学習意欲の向上のために、利用者同士が店舗を評価し合う「いいね!」ボタンや「いいね!数」ランキングなどのソーシャル情報[10]に関する機能も昨年度後期に追加し、今年度の実験から利用されている[11]。実験を活性化させた効果をもたらしている。ただし、ソーシャル機能の実験結果に関する解析は行っていない。

本論文では、作業履歴を活用し、各種のソーシャルデータ項目に対する因子分析を行うことにより、「被アクセス関係」、「購入・投票関係」、「自己アクセス関係」の3種類に要約できる結果を示す。利用者への「被アクセス数」と「いいね!数」という異なる因子に属するソーシャルデータの提供が、実験を活性化させた一因という結果が得られている。さらに、Googleページランキングアルゴリズムの適用、実験結果からより教育効果向上させるための改良点や操作履歴の活用方法について報告する。

2. 実験システム

2.1 システム概要

実験システムは、「店舗管理」と「Web モール」という二つの部分に分ける。前者の「店舗管理」で、利用者は Web ショップの作成からモールへの出店までの一連の作業を行う。後者の「Web モール」では、作成された店舗における買い物作業を行う。図 1 はシステムのトップ画面を示している。画面の左にあるカテゴリ (a) を選択すれば、すでに作成されているそのカテゴリ内のショップから買い物できる。(b) には常に店舗の 4 つのランキング、「いいね!数」「純売上数」「売上金額」「販売数」をランダムに表示している。

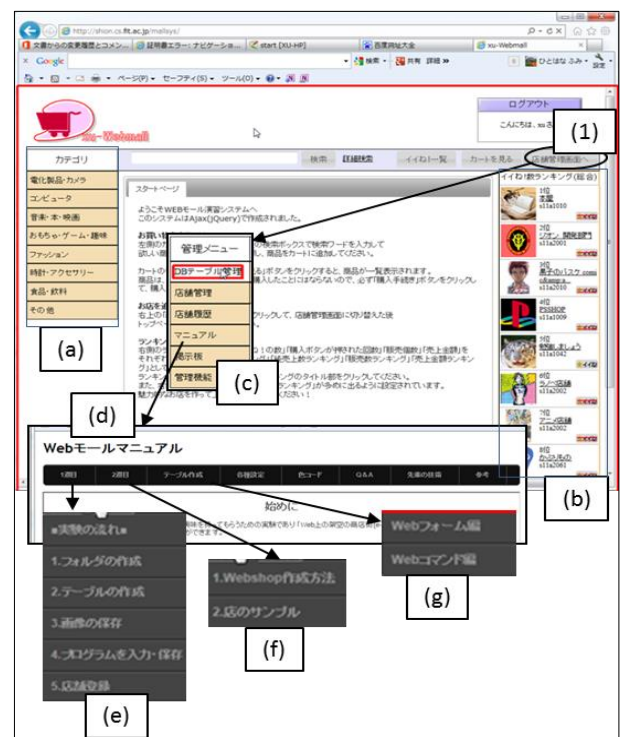


図 1 実験システムのトップ画面と遷移概要

図 1 の(1)をクリックすれば、「店舗管理画面」に切り替えられ (c)、その中の「マニュアル」項目をクリックすると、マニュアル画面 (d) に切り替えられる。実験の作業の順序と内容(e)に沿って、「1. フォルダの作成、2. テーブルの作成、3. 画像の保存、4. プログラムを入力・保存、5. 店舗登録」のように学生は Web ショップの開発を行う。作成した Web ショップを登録した後は、後半の「Web モール」部分に移り、他の利用者が作成した様々なショップを利用してもらう。商品の購入作業を体験し、ランキングの変化に注目してもらう。

2.2 ソーシャルボタンの導入

学習効果をさらに向上させるために、昨年度後期からシステムに、手軽さとポジティブ評価のみを行うという方針の元で次のようなソーシャルボタンの設置を行っている。

- ・イイね! ボタン

店舗を評価すると、イイね! ボタン(評価済) (図 2 左) になり、もう一度評価ボタンを押すと、イイね! ボタン(未評価) (図 2 右) へ変わる。



図 2 イイね! ボタン(左: 評価済、右: 未評価)

- ・店舗雰囲気評価機能の導入

より詳細に利用者の印象を集約するために、文献 [1] を参照して表 1 のような候補から 1 つを選択する店舗雰囲気評価機能も設けることにしている。

表 1 店舗雰囲気アンケート調査リスト

雰囲気名	雰囲気名
クール	今風な
明るい	古典的な
落ち着く	かわいい
楽しい	華麗
和む	洗練されている

2.3 ソーシャル情報の利用者への提供

さらなる良質店舗作成意欲の向上を図るため、各店舗のソーシャル情報を開発者に、表 2 にまとめた開発した項目でソーシャル情報をフィードバックしている。

表 2 店舗作成者に公開している情報

参照情報	解説
訪問者数	店舗へ訪れた利用者の数
アンケート結果	店舗雰囲気評価機能による評価情報
イイね! 履歴	イイね! ボタンを押した利用者情報
商品売上回数	店舗の購入ボタンが押された回数
商品売上個数	店舗での商品売上個数
商品売上金額	店舗での商品売上金額

図 3 には「イイね! 履歴」が選択された時、表示された画面が示されている。選択した人の学籍番号の最

後の一ケタを示さないことで、クリックされた履歴は伝わるがクリックした人は明かさない目的を果たしている。その他の項目については、集計情報を提供している。

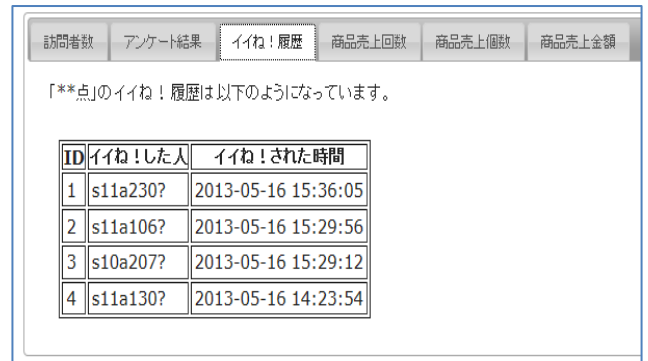


図 3 自店舗情報参照画面(イイね! 履歴選択時)

2.4 作業履歴の記録

演習システムでは、利用者の作業履歴を記録し、管理者に提供している (図 4)。さらに、より詳細な作業状況を分析できるようにするために、利用者の行った動作を全て記録するための機能(図 4 の最後の項目)を、今年度 5 月末にシステムに追加している。具体的には、いつログインしたか、どのような検索ワードで店舗を検索したか、いつ自己店舗を作成・編集・削除したかなどの情報である。

動作名	説明	実行
ショップデータ取得	システム内の全ショップデータを保存します	保存
カテゴリ情報取得	システム内のカテゴリ情報を取得します	保存
全購入履歴取得	全ユーザが購入した商品の履歴を取得します	保存
店舗アクセス履歴取得	全ユーザがアクセスした店舗の履歴を取得します	保存
イイね! 一覧取得	利用者の店舗に対するイイね! リストを取得します	保存
店舗評価定義取得	店舗の評価番号と名前を対応付けた表の取得を行います	保存
店舗評価一覧取得	管理者の店舗に対する評価の番号一覧を取得します	保存
店舗特徴定義取得	店舗の特徴番号と名前を対応付けた表の取得を行います	保存
店舗特徴一覧取得	店舗のオーナーが設定した店舗特徴一覧を取得します	保存
店舗特徴アンケート一覧取得	店舗の利用者が投票した店舗特徴アンケート一覧を取得します	保存
全ランキング情報取得	総合・各カテゴリの全ランキング情報を各 100 件記述されたデータを取得します	保存
全利用者動作情報取得	利用者のスクリプトに対してのアクセス履歴情報を取得します	保存

図 4 管理者用システム内情報取得機能の画面

2.5 関連研究

ソーシャルネットワークサービスの発達に伴い、ソーシャルデータ)の管理やマイニングに対しての研究に注目が集まっている [10]。SNS 活発化要因の解明を目指して、Web 参加者特性の解析が行われたり [3]、コミュニティの抽出 [9] が研究

されたりしている。情報推薦に関する研究も盛んに行われている[2]。

3. 作業履歴に対する解析

図 4 に示している管理者に提供しているシステムの作業履歴から、利用者の作業情報を表 3 に示しているような 8 項目に集約する前処理を、ACCESS を用いて行った。8 項目中の B, C, D, E, G という 5 項目は利用者が行われている作業情報に関する項目であり、A, F, H という 3 項目は、利用者が構築した店舗のアクセスまたは投票される情報に関する項目である（ソーシャル情報としては、図 3 に示しているように「アンケート結果」と「商品売上金額」という二つの項目のデータがまだ利用できるが、両方とも非常にバラ付きが大きかったため、今度の集計項目から外している）。

表 3 利用者ごとの作業集計項目

A	被アクセス回数（構築した店舗の被アクセスの回数）
B	アクセス数（利用者が他の店舗をアクセスした回数）
C	自己アクセス数（自分の店舗をアクセスした回数）
D	購入回数（商品を購入した回数）
E	購入商品数（購入した商品の数）
F	被購入回数（構築した店舗のアクセスされた回数）
G	イイね!投票数（他の店舗にイイね!投票した数）
H	イイね!数（構築した店舗のイイね!数）

3.1 基本集計情報

133 名の受講生の作業履歴の 8 項目の最小値、第四位分位数、中央値、平均値、第三四分位、最大値という基本統計情報を、図 5(a)に示す。なお、これは利用者の 2 週間分の作業履歴の集計である。集計作業は学期が終わった時点でを行った。2 週間の実験終わった後の放課後の時間で作業している数件のケースも含まれている。

A から H までの 8 項目に対して、学生から最も積極的な反応があったのは A（被アクセス数）という項目である。「アクセス回数があれば、見てくれた人数が分かるので、良いと思う」、「思った以上たくさんの方がアクセスしてくれた」などのような感想がレポートより寄せられてきている。最小値が 0 である項目も 8 項目中 D, E, F, G の 4 項目がある中、項目 A の最小値は 16 となっている。

集計データに対する解析を行うときに直面した問題は、数件の全体の値と大きくかけ離れた値を持つデータの存在ということである。例えば、項目 B（アクセス数）の最大値が 1000 以上であったり、項目 D（購入回数）が 600 回以上であったりということである。解析結果に差支えが出る最小限のデータのみを除くという方針の下で、図 5(b)に示した D（購入回数）と F（被購入回数）に関する制限（ $D < 300 \& F < 400$ ）を設けることによる対策に辿りついた。図 5(a)と(b)を比べると、全ての項目において、最小値と第四位分位数は同じ、最大値は大きく縮小し、中央値、平均値、第三四分位はマイナーチェンジ程度ということが分かる。

A	B	C	D
Min. : 16.00	Min. : 10.00	Min. : 3.00	Min. : 0.00
1st Qu.: 41.00	1st Qu.: 38.00	1st Qu.: 9.00	1st Qu.: 15.00
Median : 52.00	Median : 54.00	Median : 11.00	Median : 24.00
Mean : 63.52	Mean : 68.38	Mean : 14.16	Mean : 32.41
3rd Qu.: 74.00	3rd Qu.: 79.00	3rd Qu.: 15.00	3rd Qu.: 36.00
Max. : 494.00	Max. : 1092.00	Max. : 231.00	Max. : 683.00

E	F	G	H
Min. : 0.000e+00	Min. : 0.00	Min. : 0.00	Min. : 4.00
1st Qu.: 2.400e+01	1st Qu.: 10.00	1st Qu.: 5.00	1st Qu.: 12.00
Median : 1.190e+02	Median : 18.00	Median : 14.00	Median : 16.00
Mean : 3.281e+10	Mean : 29.82	Mean : 20.38	Mean : 17.22
3rd Qu.: 4.295e+09	3rd Qu.: 30.00	3rd Qu.: 26.00	3rd Qu.: 20.00
Max. : 2.890e+12	Max. : 826.00	Max. : 353.00	Max. : 65.00

(a) 133 件の全データの場合

A	B	C	D
Min. : 16.00	Min. : 10.00	Min. : 3.00	Min. : 0.00
1st Qu.: 41.00	1st Qu.: 38.00	1st Qu.: 9.00	1st Qu.: 15.00
Median : 51.00	Median : 53.00	Median : 11.00	Median : 24.00
Mean : 59.84	Mean : 59.17	Mean : 12.31	Mean : 27.27
3rd Qu.: 72.00	3rd Qu.: 78.00	3rd Qu.: 15.00	3rd Qu.: 35.00
Max. : 225.00	Max. : 146.00	Max. : 47.00	Max. : 147.00

E	F	G	H
Min. : 0.000e+00	Min. : 0.00	Min. : 0.00	Min. : 4.00
1st Qu.: 2.400e+01	1st Qu.: 9.00	1st Qu.: 5.00	1st Qu.: 12.00
Median : 1.140e+02	Median : 17.00	Median : 13.00	Median : 16.00
Mean : 1.079e+10	Mean : 22.88	Mean : 17.87	Mean : 16.87
3rd Qu.: 1.000e+06	3rd Qu.: 29.00	3rd Qu.: 25.00	3rd Qu.: 20.00
Max. : 2.710e+11	Max. : 203.00	Max. : 78.00	Max. : 46.00

(b)条件 (D<300&F<400)下での 129 件の場合

図 5 作業集計情報の基本集計情報

以降では図(b)の 129 件データについての解析を行っている。ただし、それには依然、被アクセス数が 200 以上の 1 件のかけ離れたデータが存在する。場合によっては A(被アクセス数) が 200 より小さいという条件も加えた条件 ($A < 200 \& D < 300 \& F < 400$) 下での 128 件のデータに対する解析も同時に行う。128 件のデータの場合に、A の最大値は 161.00、H の最大値が 37.00 に変わる以外は、他の項目の最大値には変化しないという状況である。

3.2 因子分析

ソーシャルデータ項目は、表 3 に示しているように 8 項目に及ぶ。因子分析は、変数の間の相関関係から共通因子を求めることで、多くの変数を少数個の共通因子にまとめて説明することを目的としている[6]。このため、表 3 中の各種の項目間の相関関係から共通因子を求めることで、多くの項目を少数の共通因子にまとめる因子分析を行った。

試行錯誤の結果で、図 6 に示された因子数が 3 である結果が得られている。第 1 因子の中で値が大きいのは、A(被アクセス回数) (0.954)、H(イイね!数) (0.882)、F(被購入回数) (0.557) の順である。この因子は利用者が構築した店舗の「被アクセス関係」情報と考えられる。

第 2 因子の中で値が大きいのは、D(購入回数) (1.002)、E(購入商品数) (0.513)、G(イイね!投票数) (0.355) の順である。第 2 因子は利用者が構築した店舗の「商品購入や投票関係」情報と考えられる。なお、第一因子と A や H 間の相関は 0.7 以上なので強い相関あり、E とは 0.4 から 0.7 間なので中程度の相関ある。第二因子と D は強い相関あり、E とは 0.4 から 0.7 間なので中程度の相関あり、G とは 0.2 から 0.4 間なので弱い相関ありという結果である。

第 3 因子の中で値が大きいのは C(自己アクセス数) (0.994)、B(アクセス数) (0.335) で、利用者のアクセス数と関わりのある項目であるので、第 3 因子は利用者が構築した店舗に対する「自己アクセス関係」情報と考えられる。なお、第三因子

とBは強い相関あり、Cとは弱い相関ありという結果である。

すなわち、表 4 に示された 8 項目は、A, F, H が「被アクセス関係」の第一因子に、D, E, G が「購入・投票関係」の第二因子に、B, C が「自己アクセス関係」の第三因子にまとめられる。A, F, H という第一因子にある利用者が構築した店舗に関する 3 項目は相関が中程度以上にあるが、利用者の行動に関する項目は購入・投票関係（第二因子）と、自己アクセス関係（第三因子）に分けられ、かつ、相関も弱くなる。

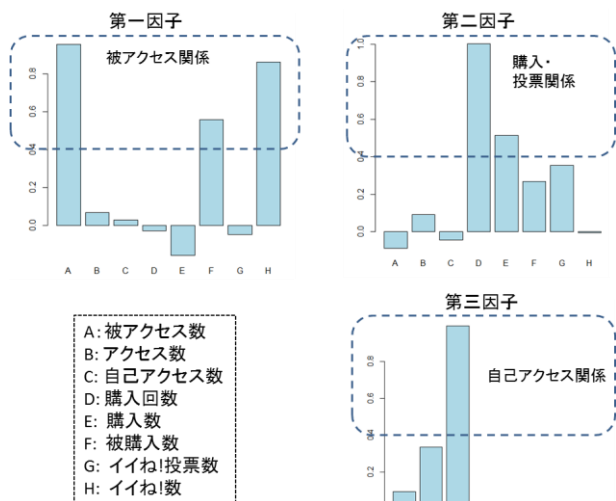


図 6 第一から第三因子の負荷量の棒グラフ

既存のシステムには購入関係のランキングと自店舗の売上に関する情報(図 3 の後半の 3 項目)しか提供していなかった。昨年度では被アクセス数(図 3 の最初の項目)を提供し、今年度では各店舗の「いいね!数」の公示と図 3 の 2 番目・3 番目の項目提供によって、実験を活性化させている。「予想以上にアクセスされているので、もっとランキング上げるように本気で作って見ようか」というように学習意欲向上のきっかけとなっているようである。

3.2 「いいね!数」に関わる解析

第一因子の全ての項目に関するデータを利用者本人への提供と H(いいね!数)の公示によって実験を活性化させたが、H(いいね!数)に関しては、次のような種類の感想が複数件寄せられている。

- ① 技術的に評価が高いとただの面白いとの区別がつかない。
- ② 公正に押すのが難しい。友達のサイトには、やはりいいね!ボタンを押してしまう傾向がある。
- ③ 完成が遅い店や偶々集計数が少ない店は、ますます注目されなくなる。

このため、本節では、H(いいね!数)について詳細に解析していく。

3.2.1 「いいね!数」に影響を与えるもの

A(被アクセス回数)と H(いいね!数)が因子分析の第一因子に属するので、その両者の関連性を調べるために、回帰分析を行った。H(いいね!数)は、A(被アクセス数)と式(a)の線形関係にある結果が得られている(図 7(a))。よりフォーカスした解析結果を得るため、128 件データに対する回帰分析も

行い、式(b)と図 7(b)で示した結果が得られている。

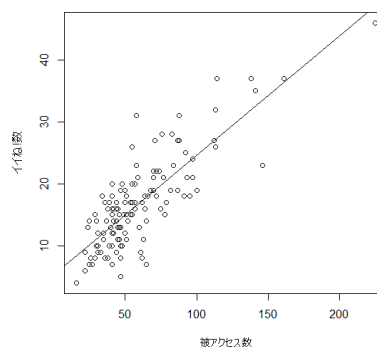
$$\text{いいね!数} = 5.33672 + 0.19269 * \text{被アクセス数} \quad (a)$$

$$\text{いいね!数} = 5.07160 + 0.19758 * \text{被アクセス数} \quad (b)$$

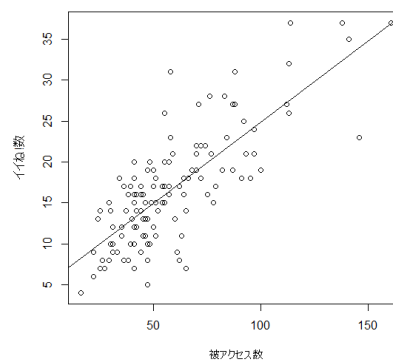
なお、両者の残差情報は次の通りである(p 値が通常よく使われている有意水準 0.05(5%)より小さく、0.001(0.1%)よりも小さい。

	Min	1Q	Median	3Q	Max
(a)	-10.8615	-3.0077	-0.0809	2.8361	14.4873
(b)	-10.9178	-2.9747	-0.1741	2.8893	14.4690

ここで注目したことは、被アクセス数が一定数(105 回)を越えた場合、「いいね!数」との残差はプラスに転じることである。すなわち、各店舗に対して被アクセス数が多ければ、「いいね!数」が必ずしも多いとは限らないが、被アクセス数が一定数を越すと、「いいね!数」は増える傾向に転じる。



(a) 129 件のデータの場合



(b) 128 件のデータの場合

図 7 (いいね!数)A(被アクセス数)の回帰直線

3.2.2 G(投票数)と B(アクセス数)間の関連

A(被アクセス数)と H(いいね!数)は店舗に関する情報である。本節では、B(アクセス数)と G(いいね!投票数)という利用者の作業情報に対する解析を行う。

図 8 に示しているように、B と G 間は線形関係になく、利用者はアクセス数もいいね!投票も行っている通常グループ(赤で示された 55 名)、投票を中心に行っている投票グループ(緑で示された 28 名)、アクセスを中心に行っているアクセスグループ(黒で示された 46 名)に分けられている。B(アクセ

ス数) と G(いいね!投票数) に対する k-means 法によるクラスタリング結果とクラスタ平均値は、次のとおりである。

	B(アクセス数)	G(いいね!投票数)	
1	87.97826	15.195652	(黒グループ)
2	39.72727	8.509091	(赤グループ)
3	50.03571	39.750000	(緑グループ)

図 8 から分かるように、アクセスしてないのに投票しているケースも見られる。すなわち、B と G の関係は利用者にもよるし、利用者と操作した店舗の作成者との関係にもよることが示されている。これで 3.2 節の感想② (いいね!ボタンの公正性) に対する裏付けにもなる。

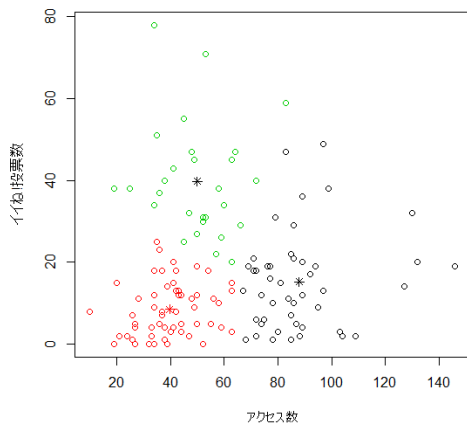


図 8 B(アクセス数)と G(いいね!投票数) によるクラスタリング結果

3.2.3 「いいね!数」と教員評価の関係

「いいね!数」の教育的効果を分析するために、各店舗の「いいね!数」と教員評価値が 5 段階に評価した「スコア」項目、および三段階に評価した「オリジナリティ」項目に対して回帰分析を行った。「いいね!数」と「スコア」の相関は 0.2657261 というあまり相関ないという結果であり、「オリジナリティ」との相関は -0.04434499 というほとんど相関なしという結果である。

3.3 編集回数を含めた後半の作業履歴の解析

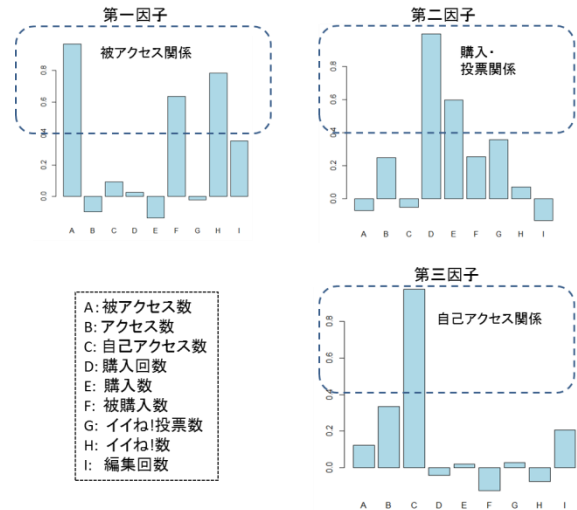
3.2 節の課題の対策を探るために、利用者の行った動作を全て記録するための機能(図 4 の最後の項目)の活用を試みた。後半 3 グループ合計 75 名に対する全履歴情報から、店舗を編集する回数という履歴情報が新たに利用できる。I(編集回数)という項目も含めた後半グループの作業履歴の基本統計情報と因子分析の結果をそれぞれ、図 9(a)(b)に示している。因子分析は、3.1 節と同じように D(購入回数)と F(被購入回数)に関する制限(D<300&F<400)を設けた結果に対して行っている。

因子分析の結果では、A(被アクセス回数)(0.967)、H(いいね!数)(0.783)、F(被購入回数)(0.636)、I(編集回数)(0.352)の順での第一因子に分けられる。D, E, G の順での第二因子とは相関なし、C, B 順の第三因子との相関も 0.208 で第一因子との相関よりも小さい。

A		B		C		D		E	
Min.	: 16.00	Min.	: 10.00	Min.	: 4.00	Min.	: 0.00	Min.	: 0.000e+00
1st Qu.	: 37.00	1st Qu.	: 38.00	1st Qu.	: 9.00	1st Qu.	: 12.00	1st Qu.	: 2.400e+01
Median	: 44.00	Median	: 47.00	Median	: 11.00	Median	: 24.00	Median	: 1.620e+02
Mean	: 47.97	Mean	: 55.82	Mean	: 12.64	Mean	: 28.01	Mean	: 1.719e+10
3rd Qu.	: 51.00	3rd Qu.	: 75.00	3rd Qu.	: 15.00	3rd Qu.	: 38.00	3rd Qu.	: 4.295e+09
Max.	: 141.00	Max.	: 146.00	Max.	: 47.00	Max.	: 93.00	Max.	: 2.710e+11

F		G		H		I	
Min.	: 0.00	Min.	: 0.00	Min.	: 4.00	Min.	: 4.00
1st Qu.	: 9.00	1st Qu.	: 5.00	1st Qu.	: 10.00	1st Qu.	: 17.00
Median	: 13.00	Median	: 12.00	Median	: 13.00	Median	: 26.00
Mean	: 19.85	Mean	: 17.68	Mean	: 13.78	Mean	: 30.92
3rd Qu.	: 24.00	3rd Qu.	: 25.00	3rd Qu.	: 17.00	3rd Qu.	: 38.00
Max.	: 203.00	Max.	: 78.00	Max.	: 37.00	Max.	: 106.00

(a) 条件(D<300&F<400)下での 73 件の基本統計情報



(b) 第一から第三因子の負荷量の棒グラフ
図 9 I(編集回数)を含めた後半のデータ

第一因子の 4 つの項目中、I 以外の項目は利用者が構築した店舗の被アクセスに関する項目であり、項目 I のみは利用者が店舗を構築する際の編集回数である。この相関はあるけど弱いという解析結果を活用するために、24 名~25 名の後半の 3 つのグループに対して、実際に考察を行った。編集回数が多ければ必ずしも店舗の出来が良いとは限らないが、確実に言える結論は、編集回数上位の 5 名以内(すなわち、上位五分の一以内)の人は、オリジナリティ高い店舗を構築していることである。

4. 解析結果による改良

本章では、3 章で浮き彫りになった課題の対策について検討する。

4.1 Google ページランキングアルゴリズムの適用

3.2 節の感想②に対しては、Google ページランキングアルゴリズム[1]の適用が考えられた。これは、Google ページランキングアルゴリズムを適用することで、すべての票は同じ重みではなく、より多く投票されたサイトを構築した利用者の一票がより重みがあるという効果があるからである。これで店舗構築より投票に熱心の利用者にけん制する役割が期待できる。

6 個のグループに分けて実験が行われているので、グループごとにおいて、投票関係のグラフの情報を利用して、Google ページランキングアルゴリズムによるランキングを求めてみた。まず作成日による差という問題に直面した。後に実施さ

れるグループの人は先に実施されたグループの作品に投票することはできるが、その逆はできないためである。このため、作成日ごとのランキング、すなわち、班ごとのランキングも提供するようにしている。

6 個のグループごとに適用後のランキングを確認した結果、ランキングに大きな変化をもたらしたが、関心度のランキングという部分が重みを増しているが、3.2 節の感想①によって指摘された技術的に評価が高いとただの面白いとの区別がつかないという課題は、依然残る。

4.2 作業履歴の編集回数の活用

感想①と③に対しては、3.3 節の結論を利用することにより、編集回数上位 5 位以内の店舗のランキングを上げる計算方法を試みた。図 10 に示しているように、ダミ店舗はグループ内のすべての店舗から投票されるが、編集回数上位の 5 位以内の店舗にのみ投票するというダミ店舗を 1 つ追加する。図の簡潔さを保つため、図 10 にはダミ店舗を始点とする枝のみを表示しており、ダミ店舗への枝は省略している。後半の 3 つのグループごとに対して適用した結果、それぞれのグループのランキングに大きく変化をもたらし、編集回数上位 5 位以内の店舗のランキングアップという目的は果たしている。

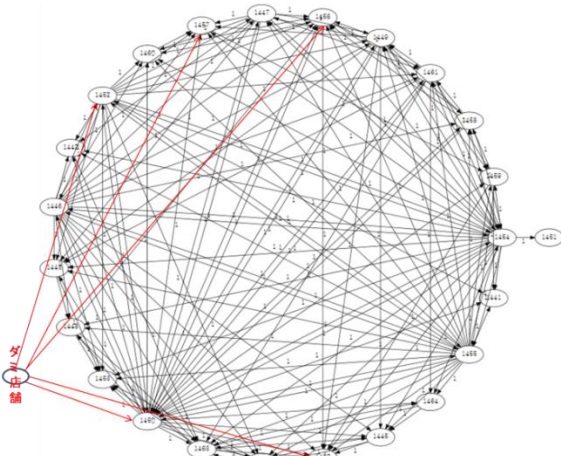


図 10 ダミ店舗を含めた投票関係グラフ

しかし、学習意欲の向上という目的に照らすと、必ずしも有効とは限らないので、この機能のシステムへの実装は見送っている。その代わりに、編集回数の上位 5 位の人の店舗を推薦機能で注目させる方法を用いることにしている。各グループに対して編集回数の平均を越した上位 5 名を「おすすめ」としてピックアップし、他の利用者に注目してもらおう。図 11 に示しているように、右側のランキングの日付を選択したときに (図 11①)、その日に作成者が 20 を超している場合、その日の編集回数上位 5 名が左下 (図 11②) 部分に表示される。すなわち、作業履歴を利用して直接にランキングを調整するのではなく、おすすめ作品として推薦することで間接的にランキングを変化させる方法を採用することにしている。これによって、3.2 節の感想③によって指摘された完成が遅い店が注目されない問題の対策にもなる。

実際に確認したところ、おすすめには出ているが、いいね!ランキングには出ない作品は、毎週 1 件や 2 件ある。データベースのテーブルに新たに列を増やしたり、SVG で視覚化を図ったりするために編集回数が多い作品がその例になる。



図 11 おすすめ作品の提供

5.まとめと今後の課題

成績評価とは切り離して、学習意欲を向上させるため一つの要素という立場で、情報工学実験科目におけるソーシャルボタン機能を導入した。積極的に実験に参加する役割を果たしている。ソーシャルデータに対する解析と考察の結果は次のとおりである。

各種のソーシャルデータ項目が、「被アクセス関係」、「購入・投票関係」、「自己アクセス関係」の 3 種類に要約できる結果が得られた。被アクセス情報に関わる項目が第一因子にまとめられ、相関は中程度である。購入と投票関係に関わる項目、利用者の行動を表す項目が。それぞれ第二因子、第三因子に分けられ、相関は弱い。本システムでは、第一因子のすべての項目を利用者に公開しているので、それと「いいね!数」の提供によって実験を活性化させている。

被アクセス回数の増加に従って「いいね!数」も線形的に増えていくが、公正性という課題も浮上した。このため、まず、アクセスしないと投票できないように、システムの投票機能の改良を行った。次に、班ごとのランキングも提供することによって、作成日によるランキング上の差が出ることを防いでいる。さらに、Google ページランキングアルゴリズムの適用と編集回数といった利用者の作業履歴情報による推薦機能の提供も行い、システムに追加している。それらによってさらなる学習意欲の向上を図りたい。ただし、現時点では、実験終了時点でのデータに対する検証しか行えなかったので、進行中の実験結果に対する解析や、学習意欲向上させるための効果的なランキングアルゴリズムの考案が、今後の課題である。さらに、商品購入状況と「いいね!数」の相関分析も今後の課題である。

参考文献

- [1] Amy N.Langville, Carl D.Meyer(著), 岩野 和生, 黒川 利明, 黒川 洋(翻訳) Google PageRank の数理 — 最強検索エンジンのランキング手法を求めて, 共立出版 2009.
- [2] Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich(著), 田中 克己, 角谷 和俊(翻訳): 情報推薦システム入門 -理論と実践- 共立出版 2012.
- [3] 金子 翔太, 小林 郁太郎: SNS コミュニティモデル化による Web 参加者特性の解明, 電子情報通信学会論文誌 D, Vol. J96-D, No. 2, pp. 280-288 2013.
- [4] 栗田 多喜夫, 加藤 俊一, 福田 郁美, 坂倉 あゆみ: 印象語による絵画データベースの検索, 情報処理学会論文誌 33(11), 1373-1383 1992.

- [5] 蔵内雄貴, 内山 俊郎, 内山 匡: マルコフ確率場を用いたソーシャルネットワークからのユーザ属性推定, 電子情報通信学会論文誌 D, 情報・システム J96-D(6) 2013.
- [6] 金 明哲: Rによるデータサイエンス - データ解析の基礎から最新手法まで, 森北出版 2007.
- [7] 筒井 亮太, 木本 博己, 松崎 大祐, 徐 海燕, 史 一華: Web モール実験システムの開発と実践, 火の国情報シンポジウム, B-4-4 2009.
- [8] 史 一華, 徐 海燕: 多様な演習システムのサービス化と統合, 教育工学研究会 (ET) 信学技報 Vol.108 No. 470 ET2008-93 2009.
- [9] 伏見 卓恭, 斉藤 和巳, 池田 哲夫, 風間 一洋: 異なる視点からのノード機能に基づくコミュニティ抽出法, 日本データベース学会論文誌, Vol. 11, No. 3, pp. 27-32 2013.
- [10] 増永 良文: ソーシャルコンピューティング入門—新しいコンピューティングパラダイムへの道標, サイエンス社 2013.
- [11] 山口 亨, 史 一華, 徐 海燕: MySQL と jQuery を利用した Web モール実験システムの開発および実験結果に対する分析, 第 120 回コンピュータと教育研究会 CE120-06 2013.