

詳細な学習履歴を活用した学習者行動の分析

中澤 真[†]
会津大学[†]
短期大学部

小泉 大城[‡]
サイバー大学[‡]
IT 総合学部

後藤 正幸^{*}
早稲田大学^{*}
創造理工学部

平澤 茂一^{**}
早稲田大学^{**}
理工学術院 総合研究所

1. はじめに

ICT の進歩とともに、学習者は教材の閲覧からテスト、ノートテイキング、コミュニケーションに至るまですべての学習活動を e-learning のシステム上でできるようになりつつある。これは学習者の活動履歴も取得しやすくなることを意味し、結果として、これらの情報を用いた学習者の進捗に応じた個別学習支援や、問題のある学生を自動抽出して教員へフィードバックする授業支援機能などの実現が期待できる。また、従来の SCORM が学習履歴の利用が LMS 任せであったのに対し、最近注目されている Tin Can API¹では汎用性をより高めた形で学習履歴を扱える枠組みを構築しようとしており、e ラーニング標準規格においても学習履歴への関心が高まっていることがわかる。

しかし、従来取り組まれてきたような LMS へのアクセス回数やファイルの閲覧回数などのシンプルな学習履歴だけでは、学習者の状態を正確に同定することは難しい。そこで本稿では、講義資料のページ単位の閲覧時間や練習問題の取り組み状況などの詳細な学習履歴を活用することで、学習者の行動特性を分析し、これを明らかにする。

2. 詳細な学習履歴活用のための実証実験

2.1 学習履歴活用環境および実験方法

まず、詳細な学習履歴を取得するために、荒本らが開発した電子教材作成支援システム[1]を用いて図1のような Web 教材を用意した。この教材を学習者が利用すると、教材名、ページ番号、閲覧開始日時、終了日時、閲覧秒数などの履歴情報を Ajax により Web サーバに自動送信されるようになっている。特に、閲覧しているか否かの判断を JavaScript の focus を用いて実現しているため、学習者が別のアプリケーションを操作したり、ブラウザのタブ切り替えで教材の閲覧を中断した場合でも、学習者の行動を正確に把握することが可能である。

実験では、PowerPoint で作成した資料をこのシステムによって Web 教材へと変換し、オープンソース LMS の Moodle²上に組み込んだ。この際、教材組み込み部分のパラメータとしてユーザ ID を設定することで、Moodle のユーザ ID も学習履歴内に記録させ、2 つの

システムにまたがる情報の紐付けを可能にした。



図 1: 電子教材のイメージ

この電子教材を用いて、会津大学短期大学部産業情報学科 1 年後期の専門科目「情報ネットワーク論」において実証実験を実施した³。電子教材をプロジェクトに投影しながら教員は授業を進め、履修学生も演習室内に設置されたそれぞれの PC 画面で自由に教材を閲覧する形態である。また、学習者の理解状況を確認するために、授業時間内に何度か練習問題に取り組ませ、これを Moodle 上のフォーラムに投稿させることにより、出題から解答までに要した時間についても履歴情報として記録されるようにした。図 2 に示したのは 2 回の授業における解答に要した累計時間のばらつきであり、この時間の差異によって十分理解している学生と不十分な学生とを区別して行動特性を評価した。

実験にあたっては、電子教材の操作性が学習に悪影響を与えていないことを確認するために実験後に学生へアンケートを実施したが、使いやすいという回答が 9 割を超え、自由記述でも従来の PDF 資料と比較して使いやすい、教員が説明している箇所との対応を取りやすいなど肯定的な意見が多く、インターフェースとしての問題がないことを確認している。

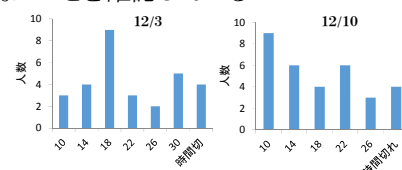


図 2: 練習問題取り組みの累計時間の分布

2.2 実験結果

取得した学習履歴から、まずシンプルに授業時間内の資料の閲覧総時間と解答時間の相関を求めた結果を図 3 に示す。相関係数は 0.12 となり、図からもほとんど関係性を見出すことはできず、単純集計では十分な分析ができないことを示している。これは授業時間外なども含めた Moodle へのアクセス回数と成績について相関が現れなかった結果と同様の傾向を示している[2]。

次に学習履歴を時系列的に捉え、教員がプロジェクトにより投影した資料のページと学生が閲覧したページ

¹ <http://tincanapi.com/>

² <https://moodle.org/>

³ 履修者数は 32 名。

のタイミングのギャップを比較し、学習の進捗が遅れている学生の特徴の抽出を試みた。平均的には教員の表示のタイミングに合わせて閲覧する学生が多いが、かなり理解している学生⁴と理解不十分な学生には異なる特徴が表れていることが明らかになった。図 4、図 5 はページごとの閲覧開始時刻をプロットしたものであり、赤丸マーカーが教員のタイミングを示している。十分理解している学生のプロットは連続していない箇所があり、飛ばし読みをしていることが伺える。また、早い段階で資料全体を閲覧しようとする行動が現れている。一方、理解が不十分な学生は 30 秒から数分教員よりも遅れることがあり、授業のペースに追いつけていないことが閲覧履歴から確認できる。

教員と学生の行動の差はページ単位の閲覧時間にも現れる。図 6 は履修学生のページ別平均閲覧時間を示したものであるが、教員の閲覧時間すなわち説明時間と大きく差が現れている箇所がいくつかある。これはインストラクショナルデザインを見直すのに有用な手がかりとなる。

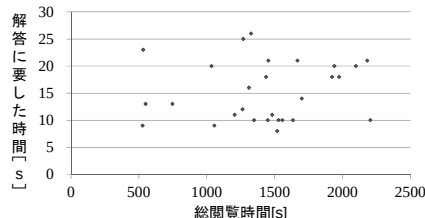


図 3: 閲覧総時間と解答時間の相関

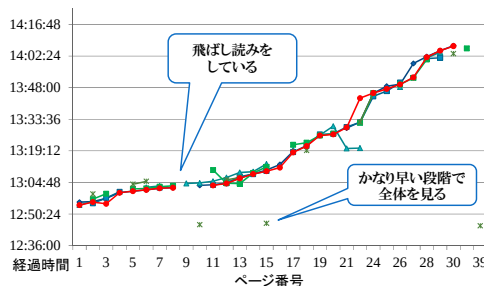


図 4: 教員と十分理解している学生の閲覧開始タイミング

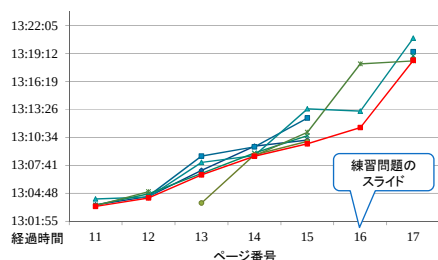


図 5: 教員と理解が不十分な学生の閲覧開始タイミング

最後に、それぞれの学生の各ページの閲覧時間を特徴量と考え、これに基づきクラスタリングした結果を図 7 に示す。今回はマンハッタン距離に基づく k-means 法により分析をし、4 つのクラスタに分類した。クラスタ 1 は他と比較して閲覧時間が短く、解答時間がやや長い学

生が集まっている。クラスタ 2 は解答時間が極めて短い学生が集まっており、問題と直接関係のない発展的内容のページの閲覧時間が長くなっている。クラスタ 3 は標準的な学生が集まっており、閲覧に関しても教員の閲覧時間と似た傾向を示している。クラスタ 4 は解答に時間がかかっている学生が集まっており、ページの閲覧時間も総じて長くなる傾向が見られる。

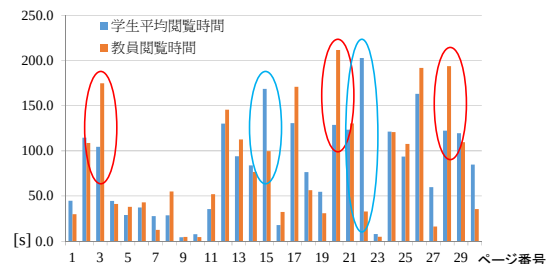


図 6: 教員の説明時間と学生閲覧時間の比較

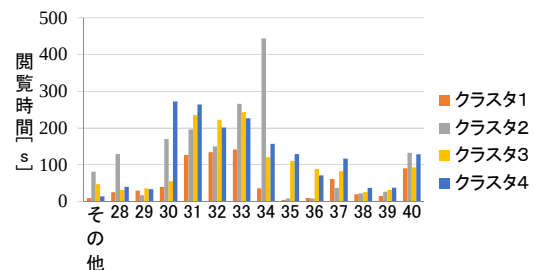


図 7: クラスタ別のページ別閲覧時間分布

3. 考察とまとめ

学生の講義資料の閲覧行動履歴を詳細に取得することによって、学習者の理解度に応じた特徴パターンが現れることを図 4・5 から確認することができた。また、これらの結果と図 6 から、教員自身の講義の進め方や作成した資料の問題発見につなげられることも確認できた。例えば、教員の説明時間に対して学生の閲覧時間が短いページの傾向として、文字が少ない図を中心としたページや発展的内容なページが多いことから、ページレイアウトや学生の興味関心を高める工夫が必要な箇所を把握することが可能である。

学生の閲覧行動に基づいたクラスタリングでは、理解度に応じた特徴を部分的に抽出することには成功したが未だ不完全な部分も多い。今後、学生の閲覧時間ではなく、ページ遷移を特徴量として分析することにより、学習者の行動特性がより明確なモデルの構築に取り組む予定である。

謝辞

本研究の一部は、会津大学短期大学部競争的研究費および独立行政法人日本学術振興会学術研究助成基金助成金基盤研究(C) 23501178の助成による。

参考文献

- [1] 荒本道隆, 小泉大城, 須子統太, 平澤茂一, “PDFファイルをベースとした電子教材作成支援システム”, 情報処理学会 第76回全国大会 講演論文集, 2014年3月.
- [2] 中澤真, 小泉大城, 石田崇, 後藤正幸, 平澤茂一, “e-learningにおける学習スタイル-協働学習と学習ログ分析”, 日本経営工学会平成25年度秋季大会, 2013年11月.

⁴ 練習問題の解答時間が短い上位者。