

# Web 利用履歴のリアルタイムモニタリングによる クラスの学習状況把握ツールの開発

渡邊 貴志<sup>†</sup> 矢吹 太郎<sup>†</sup> 佐久田博司<sup>†</sup>

<sup>†</sup> 青山学院大学 理工学研究科 理工学専攻 知能情報コース

〒 229-8558 神奈川県相模原市淵野辺 5-10-1 O 棟 526

E-mail: <sup>†</sup>c5608128@aoyama.jp, <sup>††</sup>{yabuki,sakuta}@it.aoyama.ac.jp

あらまし インターネットに接続された端末を利用する演習において, 学習者の Web 利用履歴を積極的に活用する方法を提案する. 演習室では, 学習者は検索エンジンを主体とする Web 利用によって学習を行っている. その履歴は学習者の疑問点や学習活動をリアルタイムに知るための重要な手がかりとなるはずで, それを有効に活用する手段が求められている. そこで我々は, 学習者の Web を利用した学習活動に関する履歴を取得するシステムを開発した. 本システムによって, 教師は学習者の行動をリアルタイムに把握できるようになった. 更に, 開発したシステムの試験運用によって得られたデータの視覚化に関する検討を行った. その結果をもとに, 本システムを応用することによる Web 利用中の学習者への積極的な働きかけの可能性について議論する.

キーワード e-Learning, Blended Learning, 調べ学習支援

## The realtime monitoring tool for learning state of a class using web access logs

Takashi WATANABE<sup>†</sup>, Taro YABUKI<sup>†</sup>, and Hiroshi SAKUTA<sup>†</sup>

<sup>†</sup> Course of Intelligence Information, Graduate School of Science and Engineering, Aoyama Gakuin University  
Fuchinobe 5-10-1, Sagamihara-city, Kanagawa, 229-8858 Japan

E-mail: <sup>†</sup>c5608128@aoyama.jp, <sup>††</sup>{yabuki,sakuta}@it.aoyama.ac.jp

**Abstract** We propose a method for making the most use of an internet access log in a classroom. In the classroom, students are used to access internet. Their web access log is going to be an important clue for grasping their learning status. Accordingly, a method for utilizing it is required. We develop a system for monitoring students' learning status. This system enables a teacher to grasp students' activities in real time. We test our system in three classrooms, and discuss ways to visualize the test results. In this paper, we describe outlines of our system and interactions between teachers and students who use WWW.

**Key words** e-Learning, Blended Learning, Inquiry-based Learning Support

### 1. はじめに

#### 1.1 Web を利用する Blended Learning

情報通信技術の普及に伴い, 大学をはじめとする多くの教育機関でインターネットに接続された端末が用意されるようになった. その結果, 教育現場における Web の活用が進んでいる. 特に図 1 のように, 従来行われてきた教師と学習者の対面学習に, e-Learning 環境を組み合わせる授業内容の補強を試みる Blended Learning の事例が増加している [4].

Blended Learning 環境においては, 教師は e-Learning システム上にあらかじめコンテンツや学習内容の枠組みを用意して

おくのが一般的である. しかし運用段階においては, 学習者は事前に用意された枠組みをこえて図 2 のように Web にある情報を積極的に活用しようとする.

あらかじめ学習者が Web からの調べ学習を行うことを想定して e-Learning コンテンツを用意するケースもある. Web には豊富で多様な情報があり, 調査・学習の対象として有用であるからである. 学習者は検索エンジン等を用いて Web を活用することで, 学習時に生じた疑問を教師に頼らずに解決したり, 教師からは提供されないような情報を簡単に入手することができるようになる.

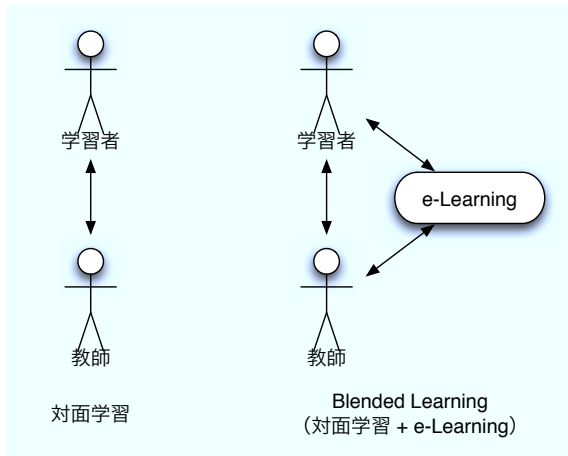


図1 対面学習と Blended Learning の比較

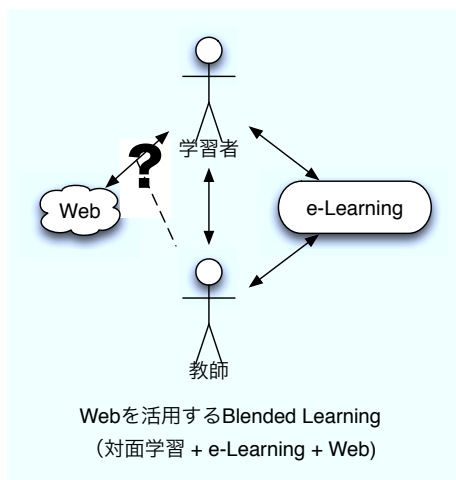


図2 Blended Learning 運用時における教師・学生・Web の関係

## 1.2 Blended Learning の問題点

Web を活用する Blended Learning の利点は次に述べる 2 点によって、教育における障害にもなり得る。

**教師側の問題** Web を利用するという学習者の行為が教師からは見えにくいため、教師は学習者の学習状況を把握しにくくなる。また、Web は教師の管理下にないため、学習者に与える影響に教師は直接関与できない。

**学習者側の問題** Web にある膨大な情報は学習を考慮して整理されているわけではないため、学習者が求める情報に到達するまでに長い時間を費やしてしまう可能性が高い。また有用な情報を見逃したり、誤った情報を参照してしまう危険がある。

## 1.3 本研究の手法

本研究では、これらの問題を解決するために以下の 3 つの仕組みを用意する。

(1) 学習者の Web 利用履歴を学習者本人や教師がリアルタイムに参照可能な仕組み

(2) 複数の学習者の Web 利用状況から知識を創発させる仕組み

(3) 学習者が閲覧する Web ページを書き換える仕組み

教師は、(1) によって Web を利用している学習者の学習状況を把握できるようになり、(3) によって、Web が学習者に与える

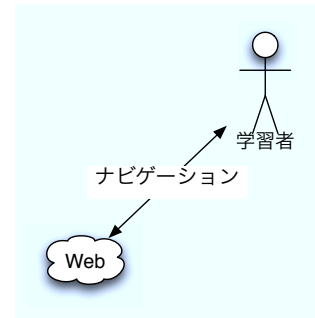


図3 学習者と Web の関係 (Web 空間のナビゲーション)

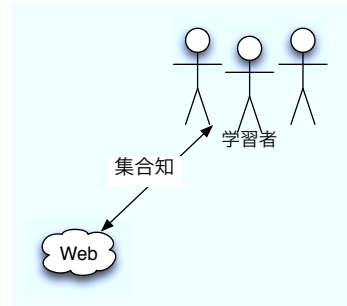


図4 学習者と Web の関係 (集合知による推薦)

影響に直接関与できるようになる。つまり、教師側の問題点が解決される。さらに (2) によって教師は学習者個人だけでなく、学習者全体の学習状況をまとめて把握することが可能になる。

学習者は、(2) によって創発した知識が (3) によって閲覧中の Web 情報に反映されるため、自分の求める情報に到達しやすくなったり、自分にとって有用な情報を見逃す危険を回避したりすることができるようになる。つまり、学習者側の問題点が解決される。

## 1.4 関連する研究

1.2 節に述べたような Web を利用する Blended Learning 環境における問題の解決は、部分的にはすでに試みられている。

### 1.4.1 Web 空間のナビゲーション

Web 空間のナビゲーションは、学習者が Web を用いて行っている調査・学習に関連する情報を整理・可視化し、図3のように学習者にナビゲーションを提示する手法である。Web ページの閲覧履歴 [6], [9] や閲覧している Web ページのリンク先 [8] に関する情報を用いたナビゲーションなどが行われている。これにより学習者は自身の状況を把握しやすくなり、同じページの閲覧を繰り返したり、学習内容と無関係な Web ページへのリンクをたどる事態を回避しやすくなる。これらの研究が想定している学習者数は 1 人であるが、Blended Learning では教室内に複数の学習者が存在するため、これらの Web 利用状況から創発する知識を利用することで、よりよいナビゲーションが実現することが期待できる。

### 1.4.2 集合知による推薦

集合知による推薦は、複数の学習者の Web 利用履歴から創発する知識を用いて、図4のように学習者が見るべき Web ページを推薦する手法である。これにより、学習者は他の学習者の情報を利用して自分の調査能力を拡張することが可能となる。高

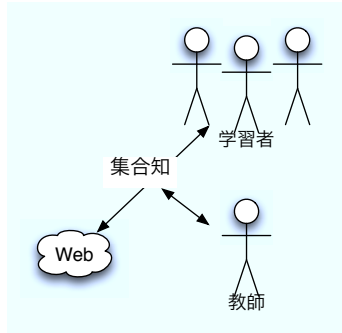


図5 学習者とWebの関係 (本研究)

須賀らの推薦システム[5]においては、対象となるWebページの数が増大になる為、システムに高い負荷がかかることが懸念される。丹羽らは、対象となるWebページを「学習者がブックマークしたもの」だけに限定することによってこの問題を回避しようとしているが[7]、ブックマークという行為自体が学習者の負担になるため、推薦のために十分なWebページが集められないことが懸念される。

## 2. 手 法

1.4節で述べた既存の手法に、学習者とWebとのインタラクションに教師が介入できる仕組みが加われば、Blended Learningの有用なツールとなることが期待される。

これを踏まえ、本研究では図5のような状況の実現をめざす。このシステム的环境下では、集合知を用いたWebナビゲーションを学習者に提供する。教師は学習者のWeb利用状況を把握し、学習に不適切と考えられるWebページを隠したり、逆に有用なWebページを推薦したりすることができるようになる。

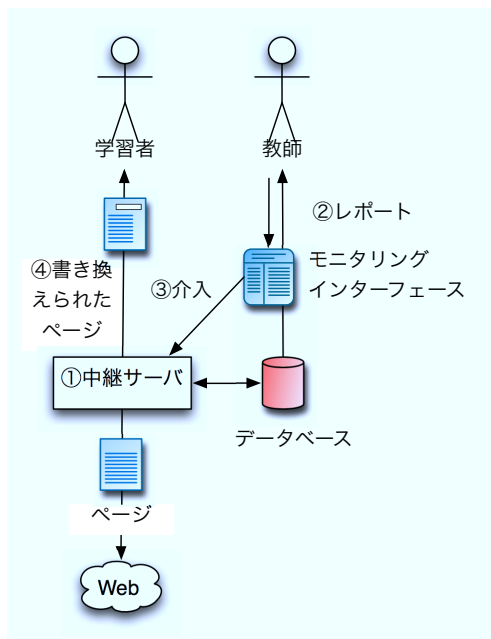


図6 システム構成

## 3. 実 装

### 3.1 システム構成

システムは図6に示されるように、中継サーバと教師用のモニタリング・インターフェースから構成される。学習者は利用する端末を中継サーバを経由してWebアクセスを行うように設定してシステムを利用する。

システムは以下のように動作する。箇条書きの番号は、図6中のそれと対応している。

(1) 中継サーバは学習者のWeb利用情報(リクエストとWebページ内容)をデータベースに格納する。さらに、学習者全員のWeb利用情報から創発される知識にもとづいて、ページに付加すべき情報があるかどうかを判断する

(2) モニタリング・インターフェースは中継サーバがデータベースに保存した情報をもとに、学習者のWeb利用状況を教師にレポートする

(3) 教師は、学習者に提示すべき情報があればそれをモニタリング・インターフェースを通して中継サーバに送る。(学習者の行動と教師の行動が同期している必要はない。教師は任意のタイミングで学習に介入できる。)

(4) 中継サーバは、集合知や教師からの情報をもとにWebページを書き換えて学習者に提示する。

### 3.2 中継サーバ

中継サーバは、広く普及しているオープンソースのProxyサーバであるSquid cache[2]をベースにPerlで開発したプラグインを追加する形式で実装を行っている。このサーバは、図7のように、次に述べる3つの機能を持つ。

- オンタイム検索検知機能
- Webアーカイブ機能
- ページ・リライト機能

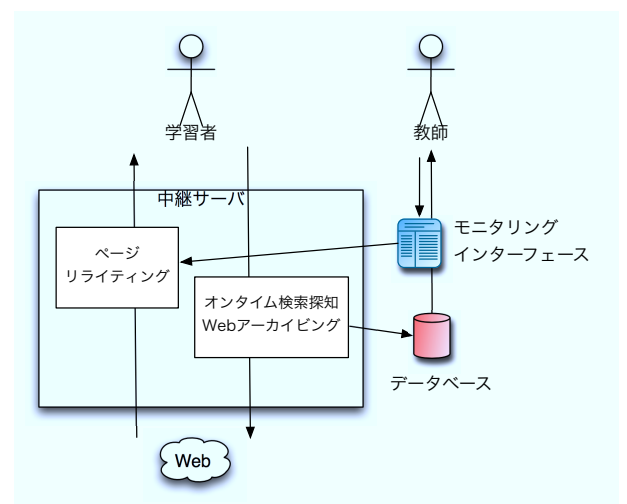


図7 中継サーバの各機能

以下、各機能を概説する。

#### 3.2.1 オンタイム検索検知機能

オンタイム検索検知機能は、検索エンジンへのリクエストを検知し、検索語を抽出する機能である。この実装は、Webページ

へのリクエストに加えて Ajax や Flash によるバックエンドのリクエストが大量に混在する状況であっても、オンタイムで分析処理をすることができる。分析処理は以下の手順で行われる。

(1) 中継サーバは学習者からリクエストを受信すると、それに対してあらかじめ用意してある検索エンジン定義データベースの情報をもとに、ドメイン名でスキャンをかける。

(2) スキャンの結果、そのリクエストが検索エンジンのドメインへのものであった場合、そのリクエストに含まれるクエリキーを抽出する。

(3) リクエスト内に有効な検索クエリキーが存在する場合、その内容の分析を行い、URL エンコードされたマルチバイト文字列が含まれる場合は適切にデコードを行う。

この機能は Google や Yahoo! など国内外あわせて 211 種類の検索エンジンを検知することができる。海外の検索エンジンへの対応には、Apache HTTP Server のログ解析用オープンソース・ソフトウェアである AWStats [1] のコードを一部利用している。国内の検索エンジンについては、独自にドメインとクエリキーに関する定義を作成した。

### 3.2.2 Web アーカイピング機能

Web アーカイピング機能は、ユーザーが閲覧した Web ページを閲覧時点での状態を保持したまま、URL や HTTP レスポンスヘッダなどの情報と共に保存する機能である。この機能は、データベース等から動的に生成される Web ページに対応するためのものである。このような Web ページには次に述べる特徴がある。

(1) 状況に応じたリクエストによってページ内容が生成される

(2) 同じリクエストであっても、時間が経過しデータベースが更新されるなどしてページ内容が変化する可能性がある

(1) のため、動的 Web ページはクローラーによって検索エンジンへ索引化されることが難しい。クローラーは、索引化の為に多種多様な状況を想定し、それぞれに応じたリクエストを生成してアクセスすることになるが、これはリクエスト先に非常に負荷をかけるためである。この結果、動的 Web ページの多くは索引化されず、一般に深層 Web ページと呼ばれる。深層 Web ページは、索引化されている Web ページの 100 倍以上存在するとされ [3]、この中に検索エンジンのみでは発見できない学習に有用な情報が存在する可能性がある。Web アーカイピング機能によって学習者が生成するリクエストのロギングを行うことで、このような深層 Web ページの把握に関する問題を解決することができる。

しかしリクエストのロギングだけでは不十分である。なぜなら (2) のために、ログに記録されているリクエストを用いても必ずしも学習者が閲覧した時と同じページ内容を取得できるとは限らないからである。Web アーカイピング機能によって Web ページの内容をリクエスト毎にテキストとしてデータベースに保存することで、時間経過によるページ内容変化の問題を回避することが可能となる。

### 3.2.3 ページ・リライティング機能

ページ・リライティング機能は、ユーザーからのリクエスト

に対する応答結果を、任意のトリガーとフィルターによって書き換える機能である。これにより、教師から学習者へのフィードバックを Web ページ内に埋め込むことが可能となる。ページ・リライティング処理は以下の手順で行われる。

(1) 中継サーバは、ある Web ページへのリクエストがトリガーによってリライティングの対象と判断された場合に、その Web ページの内容の取得先を本来の URL からフィルターへと切り替える。

(2) フィルターは、本来の URL をパラメータとして受け取り、その URL からページ内容を取得して固有の処理を行い、その結果を中継サーバへと返送する。

(3) 中継サーバは、フィルターからの返送結果、つまりフィルター固有の処理が行われたページ内容を学習者に提示する。

トリガー及びフィルター処理部は API 化されており、HTTP による入出力が可能な様々なプログラミング言語で処理を記述することができる。

### 3.3 モニタリング・インターフェース

モニタリング・インターフェースは、オンタイム検索検知機能や Web アーカイピング機能によって収集される情報を、図 6 のようにリアルタイムに表示する教師向けアプリケーションである。これにより、教師は学習者が Web を利用して行っている調査・学習の状況を把握しやすくなる。また、このインターフェースを通してページ・リライティング機能のトリガーとフィルターを編集する事が可能であり、教師は学習者と Web 間のインタラクションに対する介入を適宜行うことができる。

## 4. 運用実験

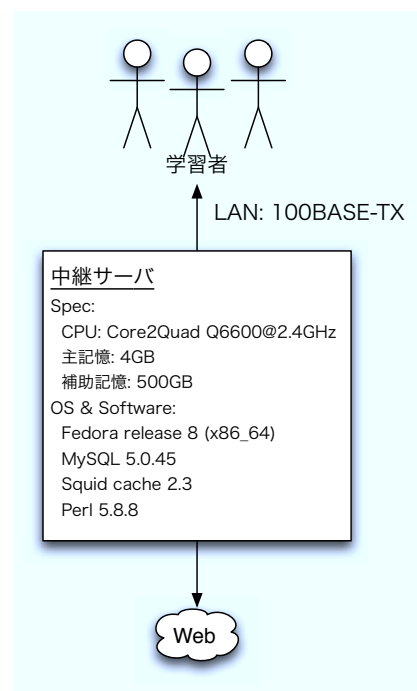


図 8 実験環境

実装したシステムに関して、基本性能の確認と実際の授業内の Web 利用情報を収集することを目的に、青山学院大学で行わ



れている次にあげる3つの授業内での運用実験を行った。

- 情報テクノロジー実験Ⅰ (Web アプリケーションを構築する実習：学習者数 42 名)
- 設計情報工学 (機械製図の実習：学習者数 112 名)
- 情報総合プログラミング実習Ⅰ (プログラミング言語でアプリケーションを構築する実習：学習者数 93 名)

実験環境は図 8 に示す通りである。

その結果、どちらの授業においてもシステムは有効に動作し、レスポンスの遅延などの Web アクセスへの障害は認められなかった。また、Linux マシンをクライアントとした負荷テストを行った結果、1 分あたり 6000 アクセス程度の処理までは中継サーバの動作に問題が無いことを確認した。

## 5. 結果と考察

運用実験によって得られたデータから、モニタリング・インターフェースを用いて教師が効果的に学生と Web の間のインタラクションに介入するため、リアルタイムにデータを視覚化する手法について検討を行った。視覚化の対象は、中継サーバによって収集される以下の2点に関するデータである。

- Web 閲覧履歴
- 検索履歴

### 5.1 Web 閲覧履歴の視覚化

3.2.2 節で述べたように、中継サーバは Web アーカイピング機能によって深層 Web ページへのリクエストを含む Web アクセスログをデータベースへと保存する。学習者が閲覧した Web ページに関する履歴について視覚化を行うことは、学習者が Web から行っている学習状況の把握に有用であると思われる。しかし収集されたアクセスログには、学習者が閲覧した Web ページ自体へのリクエストの他に、そのページが利用している CSS や画像へのリクエスト、Ajax や Flash によってバックエンドで行われるリクエストが大量に混在する。このため、前処理としてアクセスログから適切なものを抽出する必要がある。本研究では、以下に述べる条件を適用した結果を学習者の Web 閲覧履歴と定義した。

表 1 Web 閲覧履歴の抽出条件

Content-Type	text/html, text/plain
ページ内容	html タグによるマークアップが存在

この定義によって抽出された Web アクセスログについて、PHP とオープンソースのグラフ描画ソフトウェアである GraphViz を用いて、図 9 に示すような時系列による有向グラフでの視覚化を行った。

ノードは保存されたページ内容から抽出されたタイトル要素によってラベル付けが行われている。また、ノードをクリックすることで該当する URL へアクセスできる。エッジは学習者が Web ページをたどった順序を意味している。このような時系列での有向グラフ表現によって、教師は学習者が Web から行っている学習状況を俯瞰することができる。また学習者にこれを提示することで、自身の Web 閲覧履歴を直感的に理解することが可能となる。しかし、この表現は以下の点において課題が残る。

(1) 履歴数が増大するとグラフが複雑化するため、特に学習者全体への適用が難しい

(2) ノードが増えるたびに 1 から再描画が行われるためグラフの形状が変化しやすく、頻繁な更新が難しい

(1) については、履歴情報の抽象化を行うことで緩和することができるだろう。また、(2) については独自のグラフ描画の実装などを行う必要があるだろう。時系列ではなく、ページ同士のリンク関係をグラフにすることも考えられる。

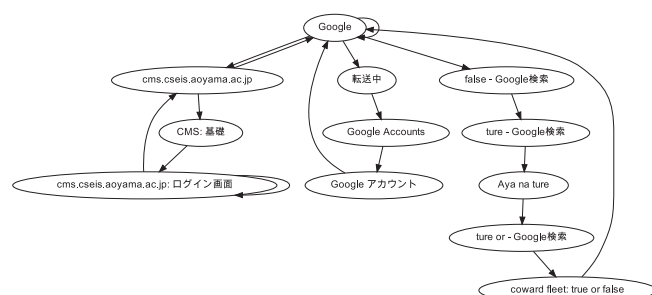


図 9 時系列による有向グラフを用いた Web 閲覧履歴の表示

### 5.2 検索履歴の視覚化

検索語は、学習者が疑問に感じている内容が学習者自身によって表現されたものである。この履歴を授業中にリアルタイムで可視化できれば、教師は授業内容に学習者の疑問への対応を反映させやすくなる。検索語は、中継サーバのオンタイム検索検知機能によってデータベースに保存されていくため、これを参照することで容易に利用することが可能である。ここでは以下の3種類の方法で視覚化し、検討した。

- テーブル
- タグクラウド
- 語順グラフ

#### 5.2.1 テーブルによる視覚化

IP	Engine	Word	Date
133.2.2.10	Bing	青山学院大学 情報	3:11:54
133.2.2.10	Google Japan	青山学院大学	3:11:27
133.2.2.10	Bing	aoyama	3:11:19
133.2.2.10	Google Japan	青学 情報	3:10:44
133.2.2.10	Google Japan	hashmap キー 取得	2009-10-26 18:10:27
133.2.2.10	Google Japan	hashmap キー 取得 api	2009-10-26 18:10:02
133.2.2.10	Google Japan	平均情報量	2009-10-26 18:09:57
133.2.2.10	Google Japan	情報量	2009-10-26 18:09:36
133.2.2.10	Google Japan	hashmap キー 取得	2009-10-26 18:09:27
133.2.2.10	Google Japan	java log 対数	2009-10-26 18:08:45
133.2.2.10	Google Japan	単語 頻度 java	2009-10-26 18:08:22
133.2.2.10	Google Japan	単語頻度 java	2009-10-26 18:08:17

図 10 テーブルによる検索状況表示

テーブルによる視覚化は、図 10 のように学習者が 1 回の検索で用いた検索句を時系列順にテーブルに並べていく表現である。このようなテーブルによる表現によって、教師は学習者の疑問をリアルタイムに把握できる。また学習者にこれを提示することで、同じ検索句の連続利用を抑え検索語の追加による検索結果の絞り込みへの誘導が期待できる。これは、同一の検索句による検索結果のページを連続で閲覧する事に対応して、テーブ

