Webシステム操作ログ保存・閲覧のための汎用的システムの開発

吉村大佑[†1]　田岡智志[†1]　渡邉敏正[†1]

現在の情報化社会のインフラとして、Web システムは必要不可欠な役割を果たしており、様々な Web システムが存在している。Web システムにおいて、システム上でユーザが行なった操作やシステムの動作をログ（動作記録）として保存しておくことは様々な点で有用である．例えば，システム管理者はログからユーザの行動を分析することによって，システムのユーザビリティ向上に資することができる．また，ユーザからの問い合わせに対しては，ログによって当該ユーザの操作，行動が把握でき，問い合わせの経緯なども分かるため，迅速に，的確に対応できる．さらに，不正アクセスの監視をすることも可能である．Web システムのログは，Apache などの Web サーバプログラム (httpd など) により保存できるが，その使用に際して情報が必ずしも十分ではない，あるいは分かりにくく分析に手数がかかる，などの指摘が多い．そのため、Web システムを作成する際に、事細かにログを保存する機能を実装することが行なわれる．しかしながら，Web システムは様々な形態が存在し，Web システムそれぞれの環境ごとにログ保存機能の実装方法が異なることから，ログ保存機能を実装するために大きな労力・費用がかかる．我々は，これらを軽減するために，システムに容易に組込むことが可能で，汎用的性を有する操作ログ保存・閲覧システムの開発を行なっており，本稿はその報告である．

Development of a Versatile System for Storing and DisplayingOperation Logs of Web Systems

Daisuke Yoshimura,[†1] Satoshi Taoka[†1]
and Toshimasa Watanabe[†1]

Web systems play an essential role, as infrastructure, in information technology of present society, and there are a wide variety of Web systems available. User operations and/or system processing of Web systems are recorded as logs in Web servers, and they are useful in maintaining Web systems. For example, a system administrator can contributes to increasing usability by catching tendency of user operations by analyzing logs; when some queries from users arrive at the system administrator, he can return quick and accurate response to them, because background of queries, and tendency of operations and actions of such users can be extracted from logs. This also helps us watch unauthorized accesses. Logs of a Web system are kept by the Web server program such as the Apache (httpd, etc.), while it is often pointed out that enough information is not always included and that listing of items is not easy to understand. This makes us spend long processing time for analysis. We try to incorporate a function to keep detailed logs as much as possible in implementation of Web systems in order to improve this situation. This means, however, that implementation of log processing heavily depends on characteristics of Web systems, and individual handling is needed. This requires much effort and high cost. For the purpose of improving such a situation, we have been developing a versatile system for storing and displaying logs such that adding this system into existing Web systems is easily done, and this is explained in this report.

1. はじめに

現在の情報化社会のインフラとして、Web システムは必要不可欠な役割を果たしており、様々な Web システムが存在している。Web システムにおいて、システム上でユーザが行なった操作やシステムの動作をログ（動作記録）として保存しておくことは，Web システムを管理，運営する上で有用である．Web システムのログとは，Web サーバの動作記録で，誰が，いつ，どこで，何をしたかを記録している．具体的には，ユーザ名，アクセス元の IP アドレス，アクセスされたページ，アクセスされた時刻などである．これらのログを利用することによって例えば，システム管理者はログからユーザの行動を分析し，システムのユーザビリティ向上に資することができる．また，ユーザからの問い合わせに対しては，ログによって当該ユーザの操作，行動が把握でき，問い合わせの経緯なども分かるため，迅速に，的確に対応できる．さらに，不正アクセスの監視をすることもできる．

Web システムのログは，Apache などの Web サーバプログラム (httpd など) により保存できるが，その使用に際して情報が必ずしも十分ではない，あるいは分かりにくく分析に手数がかかる，などの指摘が多い．そのため、システムを作成する際にはユーザがシステム上で行なった操作を再現できるようなログを保存するために，事細かにログを保存する機能を実装することが行なわれる．しかしながら，システムは様々な形態が存在し，それぞれシス

†1 広島大学大学院工学研究科
Graduate School of Engineering, Hiroshima University

テムの環境によってその機能を実装する方法が異なることから，ログ保存機能を実装するために大きな労力・費用がかかる．これらを軽減するために，我々は，システムに容易に組込むことが可能で，汎用性を有する操作ログ保存・閲覧システムの開発を行なっている．

すでにこのプロトタイプを不動産ナビゲーションシステム “賀茂ナビ”[1),2)]，広島大学大学院生物圏科学研究科における教育記録システム[3)] に導入し試運転を行なっている．

本稿では，開発・実装している本システムの概要について報告する．

2. 既存システム

操作ログの保存，閲覧，解析などを行なうシステムは大きく分けて 2 つのタイプが存在する．

1 つは，Web サーバプログラムで保存される操作ログを解析するタイプである．既存システムとしては，Visitor[5)] や WebLog Expert[6)] などが存在する．このタイプは，Web サーバに保存されているログをそのまま解析するので，システムに特別な工夫をする必要がない．そのため，導入は容易であるが，ログ自体は Web サーバが保存しているものなので，その情報だけでは不十分であったり，分かりにくいという問題点がある．

もう 1 つのタイプは，Web サーバが保存するデータとは別に，独自に各種ログを保存する機能を追加するものである．この機能をシステムに組み込む必要があるため，汎用性や組み込みの容易さなどがポイントとなる．既存システムとしては，Webjig[4)] や User Heat[7)] などが存在する．これらのシステムは，ユーザのマウスの軌跡，クリックされた座標などを記録し，それらのデータを Web ページ上に表示させることによって，ユーザの動きを可視化することができる．これらのシステムは 1 行程度プログラム変更をするだけで容易に組み込むことが可能である．しかしながら，保存されるデータ量が膨大になってしまうことや，Web ページ上に表示させないと操作内容が分かりにくい，ユーザ毎にログを区別することが難しい，などの問題点がある．また，これらはユーザビリティ向上のために利用されることが主目的となっており，システムの管理に利用することは通常想定されていない．

3. 提案システムの概要

3.1 設 計 方 針

提案システムは先ほど述べたような問題点を解消するために，以下のような方針で開発を進めている．

(1) 保存された操作ログだけから，実際の操作内容を把握できる；
(2) ユーザ毎に操作ログを区別できる；
(3) どのような環境の Web システムであっても動作する；
(4) 対象 Web システムへの組み込みが容易である；
(5) Web ブラウザ上で操作ログの検索，閲覧ができる．

3.2 システムの構成

本システムは，以下のようなプログラムから構成されている．

- **ログ送信用プログラム**
  ログを保存するサーバ（以下,「ログサーバ」とする）に対して通信を行なう JavaScript プログラム．このプログラムを対象 Web システムのページ表示部分に組み込み，ログサーバ上のログ保存・閲覧用プログラムと通信する．組み込みを容易にするために JavaScript により記述している．
- **ログ保存・閲覧用プログラム**
  ログサーバで操作ログを保存し，表示するための PHP プログラム．このプログラムは，操作名や，時間，IP アドレスなどの情報をデータベースに保存すると共に，Web ブラウザ上で操作ログを閲覧するためのプログラムである．

3.3 システムの動作概要

操作ログを保存する際，本システムは次のように動作する．

(1) ユーザが対象 Web システム上で何らかの操作を行なう (図 1(1))．
(2) ログ送信用プログラムが呼び出される (図 1(2))．
(3) ログ送信用プログラムがログサーバ上のログ保存用プログラムを呼び出す (図 1(3))．
(4) ログ保存プログラムがデータベースに操作ログを保存する (図 1(4))．

保存されたログを閲覧する場合は，ログ閲覧用プログラムがデータベースからログを読み込み表示する (図 1(5))．

3.4 操作ログ閲覧機能

本システムでは，保存された操作ログを Web ブラウザを用いて閲覧，検索することが可能になっている．通常，操作ログはサーバ内に保存されているため，そのサーバにコンソールでログインするかリモートログインして操作ログを閲覧することが多い．一方，本システムでは，Web ブラウザからの閲覧を可能にすることで，簡単に操作ログの閲覧ができるようになっている．また，ユーザの操作ログ閲覧は，システム管理者など限られた人のみが扱うべきものであるため，認証機能を実装している．

現在は，アカウント，操作名，時間を指定して検索することが可能になっている．(図 2,

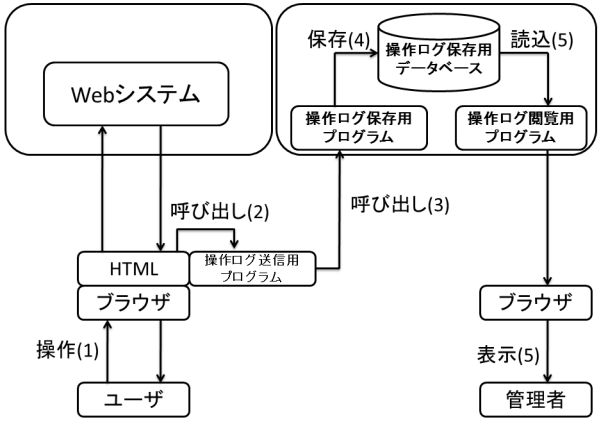

図 1　提案システムの概念図

図 2　操作ログ検索画面

3).　今後，実装が必要な機能としては以下のようなものが考えられる．

- グラフ表示など，操作ログの可視化．
  ページごとのアクセス数などをグラフ表示する．数字や文字だけで表示するよりも，視覚的に表示することによって，ユーザの操作傾向を把握しやすくなる．また，さまざまな条件で表示させることができれば，システムの性能評価に利用できるのではないかと考えている．
- ユーザ数等のリアルタイム表示．
  対象 Web システムに現在ログインしているユーザ数や，ユーザの操作をリアルタイムに表示する．これは，怪しい動きをしているユーザの特定などに利用できる．
- 通常と異なる IP アドレスやブラウザ情報を持つユーザの表示．
  操作ログには，IP アドレスや使用されたブラウザ情報などがある．これらの情報に通常では出現可能性が低いと思われるデータが含まれている場合，他人がそのユーザになりすまし，不正にアクセスしている可能性があると考えられる．そのため，通常と異なる情報を持つ操作ログを表示させることによって，不正アクセスの監視に使用できると考えている．

アカウント名:daisuke
日時:2010/01/16 16:02:15
IPアドレス:133.41.33.42
OS/ブラウザ:Mozilla/5.0 (X11; U; FreeBSD i368; en-US; rv:1.8.0.8)Gecko/20061115 Firefox/1.5.0.8
操作名:ログイン
アカウント名:daisuke
日時:2010/01/16 16:02:20
IPアドレス:133.41.33.42
OS/ブラウザ:Mozilla/5.0 (X11; U; FreeBSD i368; en-US; rv:1.8.0.8)Gecko/20061115 Firefox/1.5.0.8
操作名:新着情報ページ閲覧
アカウント名:daisuke
日時:2010/01/16 16:03:52
IPアドレス:133.41.33.42
OS/ブラウザ:Mozilla/5.0 (X11; U; FreeBSD i368; en-US; rv:1.8.0.8)Gecko/20061115 Firefox/1.5.0.8
操作名:新着情報の更新
アカウント名:daisuke
日時:2010/01/16 16:04:04
IPアドレス:133.41.33.42
OS/ブラウザ:Mozilla/5.0 (X11; U; FreeBSD i368; en-US; rv:1.8.0.8)Gecko/20061115 Firefox/1.5.0.8
操作名:ログアウト

図 3　操作ログ検索結果

3.5　システムの動作環境

操作ログを保存するログサーバに関しては，開発言語は PHP である．また，操作ログの保存先としては，データベース管理システム (RDBMS) である MySQL を使用している．RDBMS を使用することによって，保存されたデータの検索を効率的に行なうことができる．RDBMS には MySQL 以外にも様々なものが存在するが，本システムでは，単純なデータの挿入，検索のみしか行なわず，保存されるデータも短いテキストということが想定されるため，一般に検索速度が高速な MySQL を使用している．

操作ログ送信プログラムに関しては，JavaScript によって開発している．その理由としては，JavaScript はクライアント側のブラウザ上で動作するため，Web システムの構成に関わらず動作するからである．これによって，対象 Web システム構成に依存せず，操作ログ送信プログラムを組み込むことが可能となっている．

4. Web システムへの組み込み方法

対象 Web システムに，本システムを組み込む方法について説明する．ただし，セキュリティ対策を施したシステムを組み込む方法については 5 節で述べる．

ログを保存するためのデータベースを作成し，ログ送信用プログラムとログ保存・閲覧用プログラムに URL などを設定する．対象 Web システム上で操作が行なわれたときにログ送信用プログラムが呼び出されるように HTML タグの onclick 属性，onsubmit 属性などにログ送信用プログラムを呼び出す記述を指定する．このとき，操作ログとして保存する操作名を自由に設定することができる．

加えて，組み込みをより容易に行うことができるように，Web ページ上の全てのリンクやフォームなどに自動で操作ログ送信用プログラムを設定する JavaScript プログラムを実装している (図 4 参照)．これを対象 Web システムの表示部分に記入するのみで (図 4 の 6 行目)，対象 Web システムを操作すると，操作ログ送信プログラムを呼び出すように設定される．このことは実際には，ユーザの Web ブラウザが対象 Web システムから HTML データを受信した後に，onclick 属性，onsubmit 属性を必要に応じて変更することで実現している．このとき，操作名は各 HTML タグの name 属性や href 属性の値になる．

保存された操作ログを閲覧すことは，閲覧用プログラムにアクセスし，認証を行うことで可能となる．閲覧用プログラムでは簡単な検索を行なうことができ，その結果が表示される．

以上のように，本システムは，ログ送信用プログラムを対象 Web システムに組み込むだけで，その他の機能は独立したシステムとして動作する．また，対象 Web システムに一体化させて，その機能の一部として組み込むことも可能である．

```
1. <html>
2. <head>
3. <title>Sample System</title>
4. <script type="text/javascript" src="送信用のプログラムURL">
5. </script>
6. <body onload="setLog(ユーザID)">
       .
       .
7. <a href=http://hogehoge name="操作名">hogehoge</a>
       .
       .
8. <form method="post" action="fugafuga" name="操作名">
       .
       .
9. </body>
10.</html>
```

図 4 ソースコードの変更例 (太字が変更点，setLog は自動設定用の関数)

5. セキュリティ

5.1 問　題　点

本システムは，JavaScript を用いて操作ログを送信しているため，次のような問題があると考えられる．

(1) 送信されるデータを改竄することができる．

(2) ブラウザで JavaScript の機能を使わないように設定すれば，操作ログを送信させないようにできる．

(1) については，JavaScript により送信されるデータを，悪意のあるユーザが改竄することによって，操作ログの内容を書き換え，別のユーザになりすましたり，本来の操作とは違うログを送信させることができる．

(2) については，ユーザの意思により,「JavaScript の機能を使用するかどうか」が設定できる．しかも，サーバ側からこの設定を強制的に変更することはできない．そのため，ユーザが「JavaScript の機能を使用しない」と設定していると，操作ログがまったく保存されないことになる．

5.2 対　策　法

(1) についての対応策はすでに実装済みである．具体的な対策法は以下の通りである．

1. 対象 Web システム用のキーを生成し，ログサーバとそのキーを共有しておく．
2. サーバ側で，送信する操作ログ（操作名，ユーザ ID など）とキーとのハッシュ値を生成するプログラムを対象 Web システムに組み込み，ハッシュ値を生成する (図 5(1))．
3. JavaScript が操作ログを送信する際に，サーバ側で生成されたハッシュ値も同時に送信する (図 5(2))．
4. 操作ログを受け取ったログサーバは，2. と同様にハッシュ値を生成し，その操作ログの正当性を確認する (図 5(3))．

実際の組み込みでは，ハッシュ値を生成するプログラムを対象 Web システムに組み込む必要があるので，自動設定プログラムは使用できない．そのため，HTML の onclick 属性や onsubmit 属性に，操作ログ送信プログラムを呼び出すように設定する必要がある (図 6)．

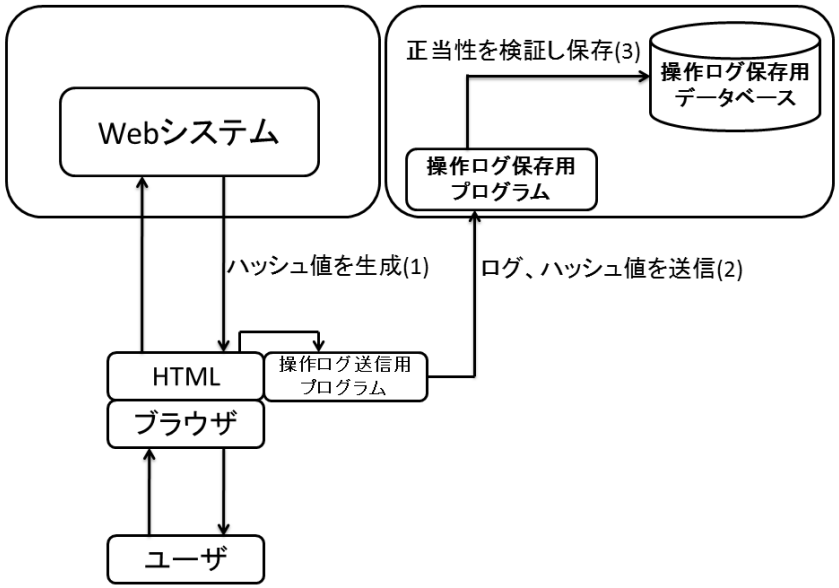

図 5 図 1 のシステムに改竄防止方策を組み込んだ改良システム

(2) についての対応策としては以下のようなものが考えられる．

- クライアントの Web ブラウザで，JavaScript 機能を有効にしないと動作しないように Web システムを変更する．具体的には，ページ遷移やフォームの送信などを JavaScript により動作するように変更することによって，ユーザが JavaScript の機能をオフにすると動作しないシステムへと変更する．

しかしながら，この方法では，本システムの組み込み易さや，汎用性が損なわれてしまうため，容易に組み込み可能で汎用性を有し，しかもこの問題点を解決するシステムに改良していくことが今後の課題である．

6. 本システムの導入事例

本システムのプロトタイプ (セキュリティ対策を行っていないもの) を組み込んだ事例を紹介する．

```
1. <html>
2. <head>
3. <title>Sample System</title>
4. <script type="text/javascript" src="送信用プログラムのURL">
5. </script>
6. <body>
      .
      .
7. <a href=http://hogehoge onclick="sendLog(ユーザーID、操作
名、ハッシュ値)">hogehoge</a>
      .
      .
8. <form method="post" action="fugafuga" onsubmit="操作ログ
sendLog(ユーザーID、操作名、ハッシュ値)">
      .
      .
9. </body>
10.</html>
```

図 6 ソースコードの変更例 (太字が変更点，sendLog は操作ログ送信用関数)

6.1 不動産ナビゲーションシステム “賀茂ナビ”[2)]

賀茂ナビは，我々の研究室で開発している東広島地区に特化した不動産ナビゲーションシステムである．Web-GIS を用いて地図による視覚的な情報掲示のもとで物件検索を行なうことができる．

6.1.1 動 作 環 境

賀茂ナビの動作環境は以下に通りである．

OS: FreeBSD

Web サーバ: Apache

開発言語: PHP

6.1.2 導 入 方 法

賀茂ナビは単純なページ遷移が行なわれる．本システムを導入するためにまず，各リンクの name 属性に，ログとして保存したときにわかりやすくなるような操作名 (リンク先のページのタイトルなど) を追加する．各ページがロードされたときに JavaScript プログラムが動作するように設定する．以上で，操作ログを保存することが可能になった[1)]．

6.2 広島大学大学院生物圏科学研究科における教育記録システム[3)]

教育記録システムは，学習・指導に関連した計画・記録・評価に関わるデータの管理・閲覧を支援する Web システムである．

表 1 本システム導入前後での応答時間の比較結果

	平均 (ms)	中央値 (ms)	最小値 (ms)	最大値 (ms)
本システム導入前	24154.9	8200.5	3058	179866
本システム導入後	24576.5	7400.5	3089	215316

6.2.1 動 作 環 境

教育記録システムの動作環境を以下に示す.

OS: FreeBSD

Web サーバ: Apache

開発言語: PHP

Web アプリケーションフレームワーク: CakePHP

6.2.2 導 入 方 法

教育記録システムでは，ページ遷移，フォームによるデータの送信，などの操作が行なわれる．本システムを導入するためにまず，各フォームの name 属性に対して，操作名を追加する．続いて，先ほどと同様に各ページがロードされたときに JavaScript プログラムが動作するように設定する．以上で，操作ログを保存することが可能になる.

7. システムの性能評価

本システムを評価するために，6.2 で説明した教育記録システム[3] に対して JMeter[8] を用いて負荷テスト行なった．テストの詳細は以下のようになっている.

- 同時アクセス数は 10 ユーザ.
- 各ユーザは以下の操作 (1)〜(5) を 10 回繰り返す.

  (1) ログイン

  (2) ページ遷移

  (3) データ送信

  (4) ページ遷移

  (5) ログアウト

負荷テストの結果を表 1 に示す．これは全ての操作が完了したときまでの所要時間を示している．この結果から，システムの導入により，平均で 400ms 程度の応答時間が長くなっていることがわかる．ただし，操作 (1)〜(5) を 10 回繰り返し，操作 (1)〜(5) それぞれで操作ログの送信，保存が行なわれるので，1 回の操作ログの送信・保存に要する平均時間は約 8ms ($= 400ms/(10\text{ 回} \times 5)$) となる．今回テストを行なった程度の規模のシステム（同時アクセス数が 10 ユーザ程度）であれば，本システムを導入してもシステムの応答時間の変化は問題にならない範囲であると考えられる.

また，今回のテストでは，全ての操作ログを正しく保存することができた．そのため，本システムの安定性に関しても特に問題はないと考えられる.

8. まとめと今後の課題

本稿では，我々が開発している「操作ログ保存・閲覧システム」について報告した．また，本システムは実際に Web システム賀茂ナビ[2], 広島大学大学院生物圏科学研究科における教育記録システム[3] に送信用プログラムを組み込み，ログ機能を利用している．さらに，負荷テストを行い，本システムの安定性，システムの応答時間に及ぼす影響を評価した.

今後の課題としては，セキュリティの強化とともに、より容易に組み込むことが出来るように改良すること、そして、より大きな規模のシステムに対して負荷テストを行うこと，などが考えられる.

参 考 文 献

1) 吉村 大佑，“Web-GIS に基づく不動産ナビゲーションシステム “賀茂ナビ”–ログと閲覧権限の管理機能強化–”, 平成 21 年度広島大学工学部第二類卒業研究論文, 2010 年

2) 濱田 友哉, 田岡 智志, 渡邉敏正, “Web-GIS に基づく不動産ナビゲーションシステム “賀茂ナビ”–情報管理機能とインターフェースの改良–”, Tech. Rep. of IEICE, OIS2007-17, pp. 19–24, 2007.

3) 入江 弘紀，吉村 大佑，田岡 智志，渡邉 敏正，“広島大学生物圏科学研究科における教育記録システムの開発”，Tech. Rep. of IEICE, LOIS2010-18, pp. 1–6, 2010.

4) 木浦 幹雄，大平 雅雄，上野 秀剛，松本 健一, “Webjig:ユーザ行動とユーザ画面の関連付けによる動的 Web サイト利用者の行動可視化システムの開発及び評価”，情報処理学会論文誌，Vol51, No.1, pp.204–215, January 2010

5) “Visitors - fast web log analyzer -”, http://www.hping.org/visitors/index_jp.php

6) “WebLog Expert - Powerful log analyzer -”, http://www.weblogexpert.com/

7) “User Heat”，http://userheat.com/

8) “JMeter”，http://jakarta.apache.org/jmeter/