

# 1ZA-6 マーケティングのための WWW サービスログ収集機構の提案

綾野雅之<sup>†</sup>, 加来田裕和<sup>\*</sup>

NTT 西日本研究開発センタ<sup>‡</sup>

## 1. はじめに

インターネット上の EC モールなど, WWW 上の BtoC サービスが活況である. これらのサービスからは, さまざまな情報を含んだログが出力されており, これを活用して効果的なマーケティングを行うことが期待される.

本稿では, WWW 上のサービスからログを収集し, マーケティングのデータとして有効に活用できる形式に変換して出力するログ収集機構を提案する.

## 2. 課題

一般的な WWW ログ解析ツールは, サーバの管理に主眼を置いている. そのためマーケティングに用いるには, 得られる情報が不十分である.

### マーケティングに必要な情報

ここでマーケティングとは, WWW サービスのログを分析し, ユーザ個人やユーザ層別に特徴的な傾向, 嗜好を見出すことである. そのためには, 以下の機能が求められる.

- 複数サービスのログを総合的に分析できる.
- ユーザの属性情報を考慮する.
- 多様な分析手法が利用できる.

### ログ解析ツールの課題

ログ解析ツールの処理概要を図 1 に示す. 処理の各段階ごとに, マーケティングでの課題を説明する.

#### (1) ファイル読込

ファイル読込, または FTP 等でログを読み込む. 大量に収集するには負荷が大きい.

WWW サーバのログ形式を予め記憶しており,

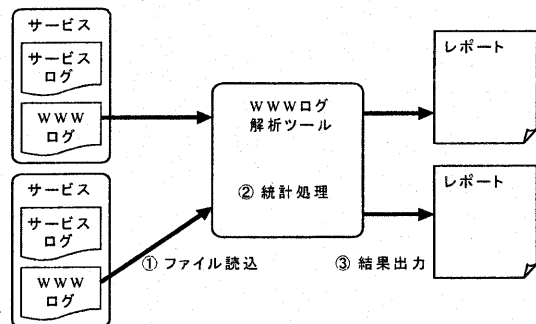


図 1 ログ解析ツール

読み込み時にカラム毎の内容を判断する. このため, 対応するログの形式が限定される.

#### (2) 統計処理

ログファイルの内容から, ページ毎, 時間毎などの単位でアクセス数を集計する. ユーザを識別しないため, 漠然とした情報しか得られない.

より詳細な分析には, ユーザを識別して他サービスのログと結合したり, 属性情報に基づいた分類が必要である.

#### (3) 結果出力

統計処理の結果をグラフや表の形式にまとめ, レポートとして出力する.

読みやすく出力される反面, 定型的内容に限られるので, データマイニングのような手法を使った分析が行えない. また, 他サービスのパラメータとして入力するのは困難である.

## 3. ログ収集機構の提案

前節で述べた課題を解決する, ログ収集機構を提案する. 図 2 がその概要である.

### (1) ログ収集

1. 圧縮して送信する

<sup>†</sup> Masayuki Ayano, Hirokazu Kakuda

<sup>‡</sup> NTT West Research and Development Center

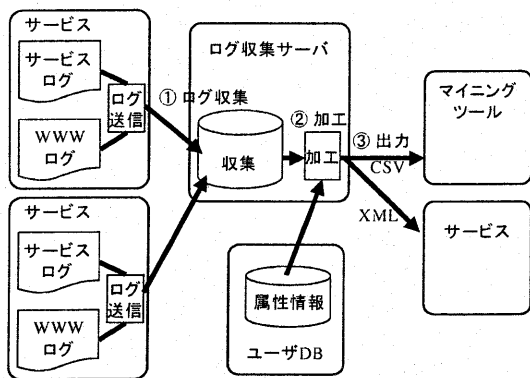


図 2 ログ収集機構

収集対象のマシンに配置したプログラムから、収集サーバにログを圧縮して送信する。集められたログは、形式を意識せずに蓄積する。

2. 分散されたログを時系列に結合する  
分割記録されたログを時系列に結合する。

(2) 加工

3. カラム毎の意味を指定する  
ログの形式を自由に設定可能とする。
4. 複数のログを結合する  
蓄積された複数のログを、ユーザ ID をキーとして結合できる。
5. ユーザ属性情報と結合する  
別に管理されるユーザ属性情報を、ログと結合して利用できる。

(3) 出力

6. 汎用的な形式での出力  
汎用的な形式(CSV,XML)での出力を行う。
7. 他サービスへの入力  
XML 形式の出力は他のサービスのパラメータとして入力できる。

## 4. 考察

提案手法と解析ツールとの比較を表 1 に示す。

課題として挙げた点は解決できたと言える。しかし、ユーザ ID が記録されていない場合には結合が行えないなどの問題もある。

## 5. 試作内容と適用例

ここでは、提案方式を元に試作中のログ収集機構の適用例を説明する。

## 適用サービス

各種の情報提供や商品販売など 17 のサービスを提供するポータルサイトに適用する。2 種類の WWW サーバログと、9 種の独自形式のログが出力されている。

## 適用の結果と課題

各サービスは独立に構築されたものであるが、統一した機構でログを収集できる。ただ、一部のログはユーザ ID を記録するよう変更した。

分析手法は未定であるが、いくつかの統計処理ツール、データマイニングツールを検討している。

実装上の課題として、ログ収集と加工が独立して動作するため、エラー発生時の動作が不定な点がある。

## 6. おわりに

WWW サービスのログをマーケティングに用いる際の要件を検討し、ログ収集機構を提案した。

これにより、複数のサービスを提供するサイト全体のログを用いて、ユーザやユーザ層の傾向を分析することを可能とした。

収集されたログを実際に分析するには、各サービスとログの内容を理解する必要がある。そのため、まず収集段階で XML など共通の形式に変換する方法も検討したい。

表 1 解析ツールと提案手法の比較

	複数サービス	ユーザ属性	分析手法
解析ツール	独立に扱う	扱えない	決められた手法
提案手法	結合して分析	結合して分析	任意のツールを利用

	収集時の負荷	統計処理	出力
解析ツール	負荷大	アクセス数	読みやすいレポート
提案手法	低負荷	行わない	汎用的な形式