

嗜好情報を用いた XML ストリーミングメディア流通基盤の構築

植木健雄 大場 正博 中沢 実 服部 進実

金沢工業大学工学部情報工学科

e-mail:{u2k,masa,nakazawa,hattori}@infor.kanazawa-it.ac.jp

近年、ネットワークの高速化に伴い、インターネットでの情報配信手法の一つとして、ストリーミングメディアが注目を集めている。既存のインターネットコンテンツである HTML や静止画像に比べ、動画と音声による圧倒的な情報量とインパクトにより、ホビーユースはもちろん、e コマースにおいても関心が高まっている。しかし、その配信方法は従来のコンテンツと同様に、ワンソース/ワンユースの関係にある。

そこで本稿では、利用ユーザの嗜好情報に着目し、個人嗜好に基づいてパーソナル化、最適化を行い、ユーザ毎にマルチストーリーに展開する、ストリーミングメディア配信プラットフォームを構築することを提案する。

A Construction of XML Streaming Media Distribution Platform using Preference information.

TAKEO UEKI, MASAHIRO OBA, MINORU NAKAZAWA and SHIMMI HATTORI

Department of Information Engineering, Kanazawa Institute of Technology

e-mail:{u2k,masa,nakazawa,hattori}@infor.kanazawa-it.ac.jp

Streaming media is becoming attractive for information distribution on the Internet due to the introduction of high-speed network. Concern in a use of hobby or e-commerce, which include video, sound and image information is increasing abruptly, compared with the existing Internet contents. However, the distribution method is based on one source/one use, like the conventional method. Then, new streaming media distribution platform is proposed in this paper, describing multi-story development based on personalized preference information for every user.

1 はじめに

近年、インターネットが急速に普及し、Web 上には多種多様な情報が氾濫している。ユーザが自身の欲する情報を入手するには、その氾濫した情報の中から手探りで取捨選択する必要がある。また、今後インターネットのインフラストラクチ

ャの整備においてブロードバンド化が加速される方向にある。そういった中、e-コマースにおいても、プライベートにおいても注目が集まっているのが、ストリーミングデータなどの映像メディアコンテンツに代表される広帯域コンテンツである。ネットワークが高速になるにつれ、より品質の高いコンテンツをユーザは求めるであろう

が、メディアの品質を向上させるだけであれば、現行の地上波 TV をインターネットで再現することと同様で、インターネットの本質的な意味での質の向上には結びついていないといえる。

本稿では、インターネットでのストリーミングメディア配信において、時間軸に展開する個人の嗜好情報を取得、分析し、さらにメディアデータに対し、マルチメディアコンテンツに特化したメタデータを付加することにより、データの分散、共有化を促進し、配信する情報のパーソナル化と情報量の最適化を行うプラットフォームについて提案する。

2 ストリーミングメディア配信基盤

2.1 現在のメディア管理

ストリーミング配信サーバは、基本的にユーザにデータを流すことだけを目的に設計されている。よって、ストリーミングメディアの管理形態は、ストリーミング配信サーバにメディアをアップロードするだけでよい。そのため、メディアデータの所有者、製作者、ジャンル等のコンテンツ情報を特定することはできない。また、メディアデータの管理を行う場合でも、各種のコンテンツ情報がなければ不便である。

2.2 XML を用いたデータ管理

メディアデータのマルチユース化、メディア管理の機能向上を目的とし、上記のようなコンテンツ情報をメタデータとしてメディアデータに付加させる。メタデータは、XML 形式で記述される。XML はマークアップ言語であるため、データの検出、抽出が容易に行える。また、XML は、データ構造さえ決まっていれば、あらゆるアプリケーションで利用可能となる。よって、本システムのストリーミング配信で用いられる SMIL も XML 形式で記述されているため、データ同士の互換性が保たれることになる。

2.3 ストリーミング配信における改善案

現在のストリーミング配信は、ライブとオンデマンドが存在する。ライブは、インターネットのグローバル性やスケーラビリティの特性を活かした配信方法であるが、現在どのようなコンテンツを閲覧しているのかを瞬時に把握することは難しい。オンデマンドはインターネットの既存の蓄積メディア同様、情報の氾濫を促進する可能性すらある。これらの要因の一つに、従来のテキストベースのコンテンツと違い、ストリーミングメディア内に包含される情報そのものを、配信サーバや利用者が把握できないということがあげられる。

2.4 提案するメディア配信方法

現在のストリーミングメディア配信では、上述したように、利用ユーザの把握と管理、メディアデータの管理を機能的に向上させる必要がある。

そこで、ストリーミング配信に利用するデータに対して、コンテンツ情報をメタデータとして付加することで、メディアデータを統合管理し、それらのメタデータと、ユーザのインタラクティブなセッションや嗜好データを解析することにより、ストリーミング配信のための記述データを動的に生成することが可能になる。そのデータを配信サーバに渡すことができれば、ユーザの状況におけるストリーミングデータのパーソナル化と最適化が行えると考えられる。

そこで、ストリーミングメディアの配信そのものには、既存のアプリケーションを用い、独自のモジュールと連携させることで、個人嗜好情報に基づくストリーミングメディア配信システムを構築する。

3 ストリーミングメディア流通システム

3.1 プラットフォーム概要

本プラットフォームは、個人情報とコンテンツ情報を解析し、配信するストリーミングデータを動的にパーソナリ化、最適化を行うことを目的とする。プラットフォームの主な機能は次の6つである

- ストリーミングデータ配信サーバ
- Webサーバ
- ユーザセッションの監視
- 個人嗜好の解析
- メディアデータの管理
- 配信データの自動再構成

これらを用いることで、個々のユーザに対し、パーソナリ化と最適化を行う。次節にて個々の機能の説明をする。

3.2 流通システム構成

プラットフォームのシステム構成は図1のようになる。

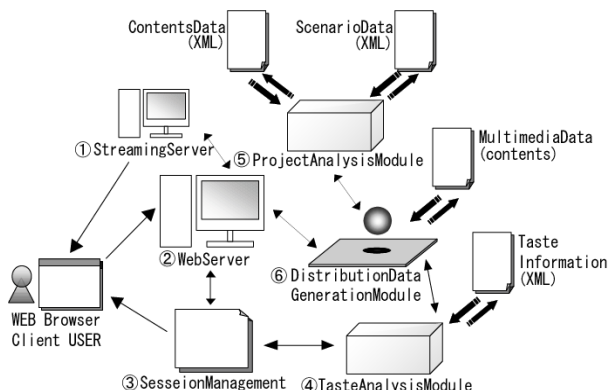


図1 配信プラットフォーム システム構成図

一連のセッションは、ユーザのWebサーバに対する、HTTPプロトコルを用いたアクセスをトリガにして開始される。ここで、本研究では、ストリーミングメディア配信における一つのセッションをプロジェクトという単位で扱う。プロジェクトは、一つもしくは複数のコンテンツにより構成され、このプロジェクトの流れをプロジェクトストーリーとして定義する。つまり、配信データのパーソナリ化とは、このストーリーが動的に変化することであり、最適化とは、プロジェクト

の大きさの調整ということになる。

Webサーバはセッションを監視しながら、リクエストに応じてプロジェクトを配信する。プロジェクト内のコンテンツには大きく分類して2種類を定義する。一つは本稿の主題になる、時間軸をもったメディアデータであるストリーミングメディア。もう一つは、時間軸を持たない、従来から用いられているHTMLや静止画像といったコンテンツである。後者に関しては、ユーザがストリーミングメディアを選択する際のナビゲーションや、ストリーミングメディアと連動して、情報の補足を行うために用いる。

個々の機能の仕様は、以下のようになる。

ストリーミングデータ配信サーバ

ストリーミングメディアを配信するためのサーバ。本プラットフォームでは、RTP、RTSPプロトコルを扱え、SMILにも対応していることから、RealNetworksが提供しているRealServerを使用する。

Webサーバ

ユーザへのナビゲーションや、ストリーミングデータと連動して情報の補足を行う。

ユーザセッション監視機能

Webサーバへアクセスしてからの、ユーザの一連の動作を監視し、セッションにおける行動の分析と、行動履歴の蓄積を行う。またセッション情報内に、プロジェクトの情報と進捗を含めることで、セッションレスなHTTP通信において、マルチストーリー性をもったストリーミング配信が可能になる。

個人嗜好解析機能

セッションから入手されるアクティブな情報と、初回アクセス時にアンケートにより取得する趣味や特徴といったプライベート情報をもとに、ユーザの嗜好や状況を解析し、プロジェクトのストーリー生成におけ

るパーソナル化のための嗜好情報を生成するモジュール。

メディアデータ管理機能

ストリーミングメディアや画像などのマルチメディアコンテンツに対して、コンテンツの内容を記述するメタデータを定義することで、メディアデータを管理する。この管理に後述するコンテンツ管理システムを用いる。また、プロジェクト中における個々のコンテンツの位置づけを定義しておくことで、ストーリーのパーソナル化を可能にする。この二つのデータを用いることで、メディアデータを総合的に管理、分散させることも可能になる。

配信データの自動再構成機能

ユーザのセッション情報、嗜好情報、メディアデータ情報をもとにして、プロジェクトのストーリーを構成する。

4 メディア管理

4.1 シーンデータ

シーンデータは、XML 形式で記述されており、各シーンで使用されるコンテンツファイルのメタデータが記述されている。1 場面で用いるコンテンツファイルへのリンク情報を記述するクリップ情報部と、その配置情報を記述するレイアウト部からなる。

4.2 シナリオデータ

1 つのプロジェクトはシナリオデータから形成されている。また、シナリオデータは XML 形式で記述されており、複数のシーンデータを含んでいる。シナリオデータは、ビューポートの定義を記述するレイアウト部と、ストーリーの流れを記述するフロー部からなる。フロー部にシーン XML のパスが記述される。

4.3 コンテンツ管理システム

本システムでは、メディアデータの管理にコンテンツ管理システムを利用する。コンテンツ管理システムは、一般的なエンタープライズアプリケーションのアーキテクチャである 3 層アーキテクチャで構成されている。図 2 は、コンテンツ管理システムのシステム構成図である。

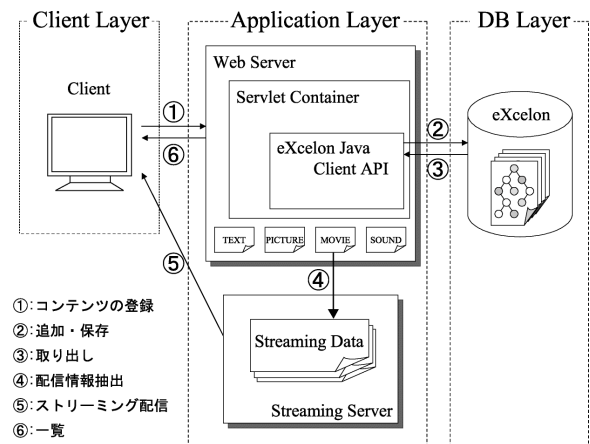


図 2 コンテンツ管理システムの構成図

クライアント層では、WEB ブラウザを用いコンテンツデータ(シナリオ XML、シーン XML、メディアデータ)の登録等を実現する。アプリケーション層では、JAVA Server Side Technology である J2EE を利用し、クライアントから要求に応じ、コンテンツデータの登録、変更等の管理処理を行う。DB 層では、JAVA との親和性が高い XML データベースシステムである eXcelon を用い、アプリケーション層で処理したデータを XML 形式で管理している。

このコンテンツ管理システムには、3 つの機能が存在する。メディアデータの管理、シナリオ XML の生成、SMIL 形式プロジェクトの生成である。メディアデータの管理は、シナリオデータ、シーンデータ、ストリーミングメディアのメタデータ、ストリーミングメディアを必要に応じて追加・変更・削除ができる。シナリオの生成は、メタ XML、シーン XML、ユーザ情報からユーザに適したシナリオ XML を生成する。プロジェク

トの生成では、シナリオ XML からストリーミングメディア用に SMIL 形式のプロジェクトを生成する。これらの処理をコンテンツ管理システムが一括管理することにより、本システムにおけるデータの同期を行い、システムを円滑に動作させる。

5 嗜好を用いたストリーミング配信方法

5.1 嗜好解析

本システムでは、ストリーミングメディアのマルチストーリー化を目的とし、ユーザ毎に異なるストリーミングメディアの配信を行う。この配信を実現するため次のような情報を用いる。情報は3種類存在し、順にユーザの初回ログイン時に取得する年齢、性別、職業、好きな色等といったプロフィール情報、ユーザのログイン時にユーザのネットワーク環境（帯域、利用端末種類）、ユーザ自身の状況（通常、急用、娯楽、調査等）をアンケート形式で取得するステータス情報、ユーザの行動履歴を解析することによってユーザの潜在的な嗜好情報を取得する履歴嗜好情報である。

ユーザのプロフィール情報、ステータス情報は、シナリオ内でのストーリーの条件分岐で用いられる。行動履歴からの嗜好情報は、ユーザの行動履歴をログデータとして管理し、ログデータを解析することにより、ユーザの潜在的な嗜好を取得する。嗜好解析手法としては嗜好空間軸と時間軸上にデータマイニングを用いることにする。

データマイニングとは、複数のデータ群を決定木、ニューラルネットワーク、相関関係等の技術を用いて分析し、データの相互関係や、ある事象におけるデータの関連性を法則化するものである。これにより、自動的にユーザの適したストリーミングメディアの配信が可能となる。

5.2 プロジェクトの推薦

本システムでは、ユーザがストリーミングメディアを閲覧する際、複数のプロジェクトを推薦することとなる。ユーザは、その推薦されたプロジェクトの中から自分が閲覧したいと思うものを複数選択する。その選択したプロジェクトのログデータを常に保存、管理しておく。ログデータの内容としては、選択プロジェクトのジャンル、プロジェクトの長さ、ストーリーの消化割合、プロジェクトで使用したコンテンツが存在する。このログデータをデータマイニングによって解析する。解析手法としては、マーケットバスケット分析を用いる。嗜好解析部の概略図は図3のようになる。

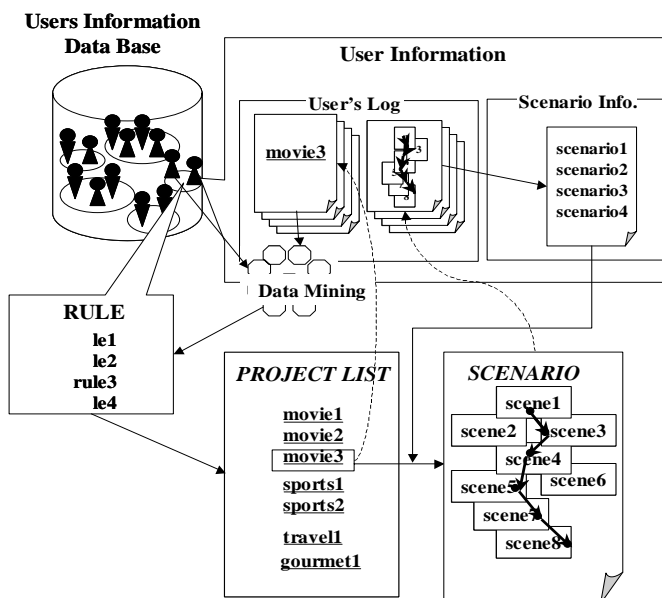


図3 嗜好解析部

マーケットバスケット分析は、ユーザの過去の履歴からどのようなプロジェクトが同時に選択されたかを統計学的に分析する。分析した結果から同時に出現したプロジェクト同士の割合を求め、その割合がある閾値を超えるとそのプロジェクト同士は、新しいルールとして、今後どちらかのプロジェクトが推薦されたら、もう一方のプロジェクトも推薦される確率が高くなる。こうすることにより、自動的にユーザの嗜好を反映したプロジェクトの推薦が可能となる。

上記のことを繰り返すうちに、必然的にプロジェクトの推薦に偏りが発生する。ユーザの管理は、サーバ側で一元管理される。そこで、嗜好が類似するユーザ同士をグループ化し嗜好ルールを共有する。

嗜好ルールの共有には、ソーシャルマイニングを利用する。ソーシャルマイニングを用いることで、ユーザ同士の嗜好が類似している場合、一方のユーザが成り立っているルールは、他方のユーザにも当てはまるという法則を導くことができる。これは、嗜好が類似するユーザ同士において、その中でルールのパターンも類似する場合、お互いがないルールを補い合う。これによって、カテゴリ化されたユーザ中における嗜好の傾向を分析し、補間しあうことで推薦の偏りを軽減するとともに、不必要な情報を抑制する。

5.3 シナリオの分岐

シナリオ分岐では、ユーザ毎に違った選択項目が設定される。これは、ユーザのプロフィール情報とステータス情報により変化する。プロフィール情報により、情報に制限をかけ、ステータス情報で、ユーザの現時点での情報が把握できるため、シーンデータ、メディアデータから時間情報を取得し、最短シナリオから最長シナリオまで生成することが可能になる。

6 おわりに

本研究では、メディアデータにメタデータを付加することで、ストリーミングメディア流通におけるメディア管理の便宜化をはかった。

また、個人嗜好を用いることで、ストリーミング配信をインタラクティブ、マルチストーリーに展開することが可能になった。ストーリーについては、ただランダムにストーリーを変更するのではなく、ユーザの嗜好を解析することにより、そのユーザに適したストーリーを展開

させた。これにより、従来の配信手法より、よりインタラクティブで表現力豊かな情報配信が可能となった。

謝辞

本研究は、通信・放送機構の地域提案型研究開発制度の支援を受けて実施された。ここに記して謝意を表す。

参考文献

- [1] W3C Thierry Michel (tmichel@w3.org) Synchronized Multimedia, 2001/02/17 <http://www.w3.org/AudioVideo/>
- [2] W3C Dan Connolly ,Extensible Markup Language (XML), 2001/02/12 <http://www.w3.org/XML/>
- [3] W3C Philippe Le Hégarret, Document Object Model (DOM), 2001/02/09 <http://www.w3.org/DOM/>
- [4] Michael J.A.Berry and Gordon Linoff , Datamining Techniques For marketing , Sales and Customer support, 1999/9/24
- [5] J.P.Bigus , Data Mining with Neural networks,1997/12/15
- [6] Marty Hall, Core Servlets and JavaServer Pages(JSP),2000
- [7] Masahiro Oba, Minoru Nakazawa, Shimmi Hattori, A proposal of Streaming Media Distribution System Based on personalized Taste information