

## 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の提案

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:{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について提案する。

### **A proposal of Streaming Media Distribution System Based on personalized Taste information.**

MASAHIRO OBA, MINORU NAKAZAWA and SHIMMI HATTORI

Department of Information Engineering, Kanazawa Institute of Technology

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

Streaming media attract attention as the technique of the information distribution on the Internet due to the introduction of high-speed network. In a use of hobby or e-commerce, which include video, sound and image and soon, concern is increasing by the overwhelming amount of information and the overwhelming impact, 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 taste 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

### 2.1 現在のストリーミング配信手段

現在、インターネット上で用いられているストリーミングメディア配信は、リアルタイムに配信するライブと、蓄積されたデータを閲覧するオンデマンドの2種類がある。サーバとビューアアプリケーション間のデータ通信には、マルチメディアデータストリームのパケットフォーマットである RTP や、その QoS を管理するために用いる RTCP が、配信データ制御にはクライアント/サーバ型のマルチメディアプレゼンテーション制御プロトコルである、RTSP などがプロトコルとして標準化されている。これらは、IP ネットワークを用いてストリーミングデータを配信する上での効率の向上と、ストリーミングデータの配送を制御するために用いられている。これは、IP ネットワーク網の帯域の狭さに伴う、非同期性や不安定さを補う目的が多くを占めており、これらの技術を用いることで、インターネット上でのストリーミングメディア配信が確立されているといえる。配信データのプレゼンテーション的な制御に関しては、W3C によって勧告された、同期マルチメディア統合言語 SMIL<sup>[4]</sup>により表現の時間的な挙動やスクリーン上のレイアウト、ハイパーリンクとマルチメディアオブジェクトの結び付けが記述でき、さらに配信サーバ独自の仕様を用いることで、静止面の制御やテロップの挿入などプレゼンテーション方法を定義することも可能である。

### 2.2 配信方法の改善点

現在のストリーミング配信は、前述のサーバの

配信機能から推察できるように、基本的にはデータ特性と IP ネットワーク特性を考慮した配信を重視しており、既存の地上波 TV 放送をインターネットで再現しているような状況といえる。ライブに関しては、インターネットのグローバル性やスケーラビリティを考えれば、特性を活かした配信方法といえるが、オンデマンドになるとインターネット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。

### 2.3 提案するストリーミングメディア配信方法

ストリーミング配信用サーバは、基本的にユーザにデータを流すことを目的に設計されている。したがって、上記の点を改善するには、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。

そこで本システムでは、ストリーミング配信方法には従来製品を利用し、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配信システム

### 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ル化、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

- ① ストリーミングデータ配信サーバ
- ② Webサーバ
- ③ ユーザセッションの監視
- ④ 個人嗜好の解析
- ⑤ メディアデータの管理
- ⑥ 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 配信システム構成

プラットフォームのシステム構成は図1のようになる。

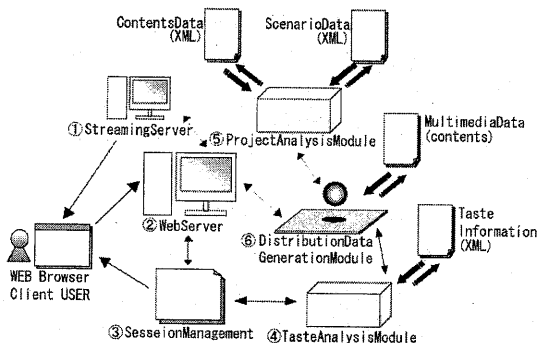


図1 配信プラットフォーム システム構成図

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。

- ① 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。
- ② 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。
- ③ 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。
- ④ 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リ化のための嗜好情報を生成するモジュール。
- ⑤ 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。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。

#### ⑥ 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. マルチストーリーの生成

#### 4.1 ユーザ嗜好解析

プロジェクトを生成する際、パーソナル化と最適化を行うためにユーザの嗜好解析を行う。これには、ユーザの現状を把握し、プロジェクトの質や、量的な最適化を行うことと、ユーザの行動履歴や嗜好情報より、プロジェクトで利用するメディアデータの選択や、ナビゲーションのパーソナル化によるユーザ支援を行うという目的がある。ナビゲーションには、ユーザがプロジェクトを選択する際と、プロジェクト内のストーリー分岐の2種類があり、プロジェクト選択の場合であれば、ユーザの情報探索をサポートし、取捨選択を効率よく行える環境提供を目的とし、ストーリー分岐時であれば、分岐数や分岐点の制限などにより、プロジェクト全体の情報量と時間の調整を行う。

ユーザの嗜好情報として、3種類のデータを用意する。

##### 1. 現在の状況（アクティブアンケート）

ユーザの現在のネットワーク接続状況（帯域、利用端末種類）や、ユーザ自身の状況（通常、急用、調査、娯楽など）、プロジェクトに対しての要求（全情報、要点だけ、動画重視、情報重視など）をWebサーバログイン時にアンケート形式で取得する。これらの情報により、プロジェクトの長さ、ストーリー分岐の数、使用メディアデータ、に変化をもたらす。

##### 2. 行動履歴

ユーザの行動を履歴として蓄積し、ナビゲーションや同プロジェクト閲覧時の展開に変化をもたらす。履歴情報として、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などの情報を取得する。この情報を利用することにより、過去の行動からユーザの嗜好を解析し、ナビゲーションのパーソナル化において反映させ、情報選択の支援を行う。

##### 3. プライベート情報

初回ログイン時にアンケートによりプライベート情報を取得する。項目としては、性別や年齢、職業といったステータスと、趣味や、好きな色、好きな雰囲気といった感性的な概念を取得する。このプライベート情報はナビゲーションを特化するパーソナル化と、一般化する最適化の2側面で利用する。

プライベート情報を用いることで、ナビゲーションをパーソナル化し、利用者が好むプロジェクトを紹介することや、プロジェクト内容に嗜好を反映することが可能になる。そのためには、プロジェクト情報や、メディアデータとプライベート情報との間でのマッチングを行う必要がある。ここで、全ユーザが個別にマッチング情報を持つとなると、コンテンツの数とユーザ数に比例して、膨大な情報量が必要になる。そこで、ユーザのプライベート情報をルール化し、プロジェクトやメディアデータとマッチングすることで、処理効率と情報量を減少させる。プライベート情報のルール化には、マーケットバスケット分析法<sup>4)</sup>を用いる。これは、統計と確率をもとに、事実の蓄積から傾向をルール化する手法である。つまり、利用者と似た状況と感性を持つユーザがこれまでにとった行動履歴が、プロジェクトやメディアデータ群とマッチングするためのルールが生成されるのである。これによ

って、無秩序に増加するコンテンツをフィルタリングすることが可能になる。

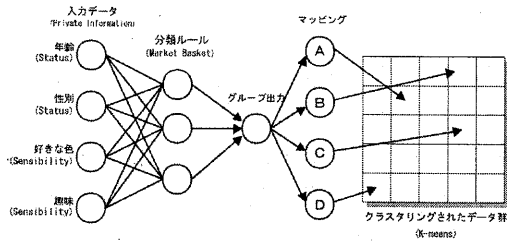


図2 嗜好とメディアデータのマッピング

情報をパーソナル化と言うことは、そのユーザに特化するということである。上記のようにユーザに特化することで、嗜好を反映したナビゲーションが可能になり、ユーザが好むプロジェクトを提供することが可能になる。しかし反面、ユーザの習慣によっては、偏ったナビゲーションをする可能性がある。そこで、これらのプライベートデータをカテゴライズすることで、利用ユーザに類似したユーザ群を見つけ出し、そのユーザ群の履歴の傾向をもとにナビゲーションや同プロジェクト閲覧時の展開に変化をもたらす。この変化によって提供する情報の偏りを緩和させる。カテゴライズには、クラスタ分析である K-means 法<sup>[4]</sup>と、リンク分析であるグラフ理論<sup>[4]</sup>を用いている。K-means 法は、分類対象の相関関係によりカテゴリを求める手法であり、この手法にプライベート情報を入力値として与えることで、ユーザをステータスと感性から分類する。ここで、K-means 法では、入力する項目の要素間に対して数値的な相関関係が必要である。そこでプライベート情報のある項目の要素間に、数値的な関連付けを行うためにグラフ理論を用いる。グラフ理論とは、各種の関係やリンクに関する理論である。グラフを用いることで、ステータスや感性というものの関係に相関関係を与え、K-means への入力値を数値化している。これらの手法によってカテゴライズされたユーザ群と利用ユーザを結び付け、全体の履歴データから利用ユーザに近い嗜好をもったユ

ーザだけを抽出して解析することが可能になる。これにより、個人に特化しすぎることによる興味範囲の極端な局所化を防止する。

これらの嗜好情報を XML<sup>[2]</sup>形式で保存し、ユーザごとに取り扱う。XML を利用することにより、機能間でのインターフェースが統一されるというメリットと、XML はオブジェクティブな仕様のため、例えば、コンテンツ情報と履歴とをリンクして解析するといったことが可能になる<sup>[3]</sup>。

#### 4.2 メディアデータ解析

嗜好解析によって得られた情報から、嗜好を反映したプロジェクトストーリーを生成するためには、プロジェクト配信者の定義する構成と、メディアデータを把握する機構が必要である。そこで、配信プロジェクトに対して1つのシナリオデータと、メディアデータに対して2つのメタデータを付加する。

##### 1. シナリオデータ

ストリーミング情報配信に双方向性、マルチストーリー性を持たせるため、プロジェクト全体のシナリオを配信者が定義する。このシナリオデータを主軸にして、利用者によるマルチストーリー性を展開していく。シナリオデータは、サブシナリオデータの集合と制御点で定義される。サブシナリオデータとは、ストリーミングメディアにより構成されるシーンや、静止画、テキスト情報で構成され、サブシナリオ毎に意味を持ち、その中で用いられるメディアデータの重要度等を定義する。制御点とはマルチシナリオを構成する上で必要な、分岐やシナリオの終点等を定義するものである。

##### 2. メタメディアデータ

サブシナリオに含まれるメディアデータ毎に、メディアデータ自体のサイズや長さ、ファイル形式や、表現するジャンルや意味、構成を記述する。このメタデータは、シナリオに特化することない汎用的なもので、メディアデータの再利用や分散化のための情報も記述される。これにより、メディアデータを

ストーリー配信と切り離し、独立したデータ群として扱える様になる。

### 3. メタスクリプトデータ

メタスクリプトデータは、シナリオデータに従属し、シナリオに特化したメディアデータの役割を記述する。シナリオデータにおける、メディアデータの位置づけを記述することにより、プロジェクトストーリーにおいて、重要な情報が欠落することを防止し、自動生成されたストーリーが無意味なものにならないように調整を行うための情報である。

これらの情報も、嗜好情報と同様に XML 形式にすることにより、お互いのリンクを可能にし、プロジェクトに特化するデータと、そうではないデータとの連結を可能にする。

#### 4.3 プロジェクトストーリー生成

配信データの最適化とパーソナル化は、ストーリーミングサーバに対する、プレゼンテーションデータ (SMIL) の発行によって実現する (図 3)。

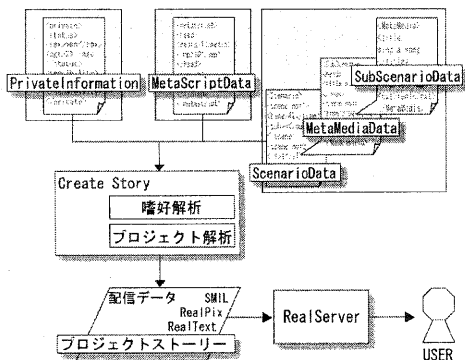


図 3 プロジェクトストーリー生成

セッション情報をもとに利用ユーザを特定し、嗜好情報と閲覧プロジェクトのメタデータ群をもとに解析を行う。解析のステップは以下のようになる。

- ・ シナリオデータによるオリジナルストーリーの展開
- ・ 利用者の状況からプロジェクト構成を仮決定
- ・ 仮決定した構成に、嗜好情報を反映させ、プロジェクト構成を決定

- ・ シナリオデータをプロジェクト構成とメタスクリプトデータをもとに再構成
- ・ 再構成されたシナリオデータに嗜好解析結果を反映させプロジェクトストーリーを生成
- ・ プロジェクトストーリーをプレゼンテーションデータに変換

こうして生成されたプレゼンテーションデータに従ってストーリーミングメディア配信を行うことで、嗜好情報に基づき、パーソナル化と最適化を可能にしている。

## 5. まとめ

本稿では、インターネットでのストーリーミングメディア配信において、個人の嗜好情報を用いて、利用ユーザにおけるパーソナル化、最適化を行い、動的にストーリーを再構成することで、ストーリーミングメディア配信をマルチストーリー化するプラットフォームについて提案した。今後は、マルチストーリー生成におけるオーサリング手法について検討していく。

## 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