

人の認知機能モデルを用いた MPEG DASH における  
ビットレート制御手法  
Rate control method of MPEG DASH based on a human cognitive model

小南 大智<sup>†</sup> 岩本 真尚<sup>‡</sup> 大歳 達也<sup>†</sup> 村田 正幸<sup>†</sup>  
Daichi Kominami Masayoshi Iwamoto Tatsuya Otoshi Masayuki Murata

### 1. はじめに

近年、動画ストリーミングサービスを提供する Over-The-Top プロバイダーが多数出現しており、モバイルデバイスを用いて動画を視聴するユーザの数が急激に増加している。モバイルトラフィックが急増することで、安定したネットワーク通信品質 (QoS; Quality of Service) の提供が困難となり、動画ストリーミングサービスの品質低下を招く可能性がある。動画アプリケーションサービスの品質を評価する際には、QoS に加えて、ユーザの体感品質 (QoE; Quality of Experience) が重要な指標として注目されている。現在、YouTube や Netflix などの動画ストリーミングサービスプロバイダーの多くは、QoS に応じて適応的にビットレートを制御する HTTP Adaptive Streaming (HAS) を採用しており、特に HAS の標準規格の一つである Dynamic Adaptive Streaming over HTTP (DASH) が普及している。

HAS における適応的ビットレート制御 (ADR) によって QoE の向上を目的としたビットレート制御を行うときには、2 点の課題がある。1 点目は QoS の変動への対応であり、もう 1 点は動画品質に対するユーザの好みの違いへの対応である。我々はこれまでに、QoS が変動する環境下においてユーザ個々の QoE を最大化することを目的とした ADR 手法を提案してきた [1]。本稿では、実機環境における提案手法の性能比較評価を行い、提案手法が QoE 向上の観点から有効であること明らかにする。

### 2. MPEG-DASH

MPEG-DASH (DASH) は、2012 年に標準化された、HTTP プロトコルを使用した動画配信技術である。DASH では、元となる一つの動画ファイルから、異なる複数のビットレートでエンコードされた複数のファイルが生成される。さらに、エンコードされた各ファイルは固定長のセグメントに分割されサーバに保存される。DASH によるストリーミングセッションが開始すると、DASH サーバは MPD (Media Presentation Description) ファイルを生成し、クライアントに提供する。MPD ファイルには、クライアントが利用可能な音声および映像のビットレートに関するメタデータが記述されている。動画を再生するために、クライアントは最初に MPD ファイルを取得し、ネットワーク状態やデバイスの種類に応じて適切なビットレートのセグメントを要求する。クライアントはセグメントごとに HTTP 要求を送信することで、再生する動画のビットレートをセグメントごとに切り替えることができる。

### 3. 提案手法の実装

DASH における動画ビットレートの選択は、クライアントに実装されている ABR アルゴリズムに基づいて行われる。一般に ABR アルゴリズムでは、クライアント端末と動画配信サーバ間の QoS の推定結果や動画再生中のプレイヤーから取得可能な情報を用いて、次にダウンロードするセグメントのビットレートを決定する。ここで、モバイルネットワークにおいてはトラフィック変動が激しいこと、ユーザが移動することなど、様々な要因によって QoS が変動する。変動によって QoS の推定が誤ると、ABR アルゴリズムが設計者の期待した動作を行わない可能性が生じる。例えば、ビットレートの切り替えを頻繁に行う、実際の QoS に対して不適切なビットレートを選択する、などが考えられ、これらはユーザの QoE の低下につながることを報告されている [2]。また、ユーザの QoE 向上を行うためには、動画品質に対する好みユーザごとに異なることから、ユーザごとに異なるビットレートの選択アルゴリズムが必要となる。動画品質に対する好みの例としては、高い画質を好むユーザ、動画再生が途中停止されないことに重きをおくユーザなどが考えられる。

本稿では、我々がこれまでに提案した、観測情報から QoS を適切に認知し、ユーザの好みに合ったビットレートを選択する手法を DASH クライアントに実装し、評価を行う。

#### 3.1 人の認知機能モデルを用いた ADR の実装

観測情報から QoS を適切に認知し、適切なビットレートを決定するための方法として、人の認知・意思決定機構に着目した。人間の脳においては、様々な感覚器からの入力情報を脳の中にある記憶と比較して意思決定を行う、トップダウン型の認知・意思決定機構が存在することが知られている [3]。この一連の情報処理がベイズ推論に基づく数学的なモデルにより説明可能であることが文献 [3] では述べられている。このモデルでは、観測情報が不確実であっても、情報の観測を繰り返し行うことで、観測した情報があらかじめ記憶した情報のいずれかと一致するのか、いずれとも一致しないのかを、逐次的に判断することが可能である。このモデルを工学的に応用することで、観測値の変動する状況においても、安定した制御が実現できると考えている。

##### 3.1.1 人の認知機能モデル

Bayesian attractor model (BaM) [3] は、人の情報知覚と弁別を、ベイズ推論の枠組みを用いてモデル化したものである。BaM は大きく三つの要素から構成される。すなわち、情報知覚、情報弁別、意思決定である。

情報知覚のモデル化において、BaM では観測対象情報を一つの特徴量としてベクトルの形式で記述する。この情報

<sup>†</sup> 大阪大学 大学院経済学研究科 Graduate School of Economics, Osaka University

<sup>‡</sup> 大阪大学 大学院情報科学研究科 Graduate School of Information Science and Technology, Osaka University

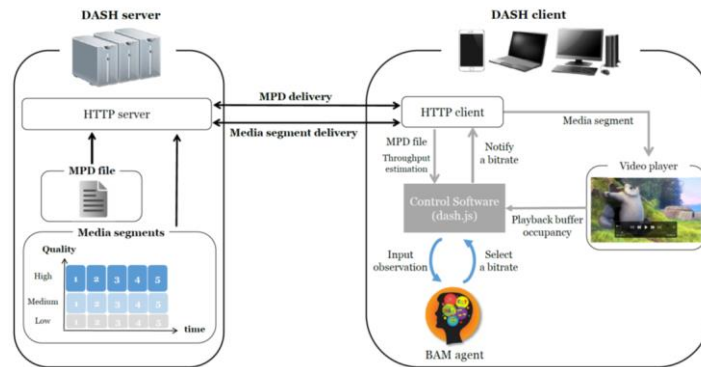


図1 実装したDASHサーバクライアント

の知覚は一定周期で行われる。情報弁別では、知覚した情報を、あらかじめ記憶しておいた情報と照らし合わせ、どれと一致するか弁別を行う。記憶する情報は、入力と同じ次元の特徴量ベクトルの形で記述される。情報弁別は  $K$  個のアトラクターを持つ非線形ダイナミクスに従って更新される状態変数  $z$  を用いてモデル化されている。ここで  $K$  は記憶している情報の個数であり、記憶とアトラクターが結びつけられている。 $z$  は、前述の非線形ダイナミクスと、知覚した情報に従って更新される。ここで、記憶する情報を表す特徴量ベクトルは、 $z$  を非線形変換することで得られるものと定義される。そのため、知覚情報から  $z$  を更新（推定）する際にバイズフィルタが用いられる。最後に意思決定モデルでは、いずれかのアトラクターに対して、あらかじめ定めたしきい値よりも  $z$  が近づいている場合に、観測した情報は、そのアトラクターに記憶した情報と一致するものとみなす。

### 3.1.2 人の認知機能モデルを用いた ADR

ADR に BaM を用いるためには、BaM を用いて認知する情報を定義し、認知結果に応じたビットレートの選択アルゴリズムを定義する必要がある。認知する情報はクライアントが観測可能な情報である必要があり、文献 [1] と同様に、本稿でも再生バッファ長と動画セグメントのダウンロード速度を用いることとする。ビットレート選択アルゴリズムも文献 [1] と同様に、ユーザの好みを表すモデルは所与である仮定のもと、認知した再生バッファ長と動画セグメントのダウンロード速度から、ユーザの好みを満たすようなビットレートのセグメントをサーバに要求する。

### 3.1.3 提案手法の実装と評価

MPEG-DASH を用いて動画を提供するサーバと、動画再生を行うクライアントを構築した (図 1)。提案手法はクライアント側で動作する。観測情報である再生バッファ長と動画セグメントのダウンロード速度は、dash.js において取得し、Web ソケットを介して BaM の動作するエージェントに送信している。BaM の動作するエージェントでは、受信した情報に従って状態変数  $z$  を更新していく。受信した情報が、あらかじめアトラクターに記憶しておいた情報と一致すると判断できた際に、ビットレート選択アルゴリズムを用いて次にダウンロードするセグメントのビットレートを決定し、それを dash.js に送信する。

提案手法の評価では、文献 [1] におけるシミュレーションで用いた設定と同様に、DASH Industry Forum の提供するベンチマーク設定を利用した。すなわち、30 秒ごとに利用可能帯域を 5.0, 4.0, 3.0, 2.0, 1.5, 2.0, 3.0, 4.0, 5.0 Mbps と変化させ、5 分間の動画を再生する。ただし、ノイズとして平均 0、分散が利用可能帯域の 10~30% の正規分布に従う乱数を追加する。この利用可能帯域の設定は linux の tc コマンドによって擬似的に生成した。評価指標として、平均ビットレート、ビットレートの変動の時間平均、およびリバッファリング（動画停止）時間を評価した。提案手法はユーザの好みに応じてビットレートを選択する機能を持つため、DASH が標準で使用している ADR と比較して、ユーザの好みに応じたビットレートの選ばれ方が実現できていることが実機においても確認できた。また、リバッファリング時間については、本評価環境では 0 秒と、動画を停止することなく、適切なビットレートが選ばれることが示された。

## 4. おわりに

本稿では、これまでに我々が提案した、人の認知機能モデルを用いたビットレート制御手法の実装および評価結果を示した。DASH の標準で使用されている ADR と比較して、提案方式が実機においてもユーザの QoE を向上できる可能性を示した。今後は、モバイル通信環境下における評価を行い、手法の有効性を確認する。

### 謝辞

This research work was supported by JSPS KAKENHI Grant Numbers 18H04096 and the Ministry of Internal Affairs.

### 参考文献

- [1] 岩本真尚, 大歳達也, 小南大智, 村田正幸, “動画配信サービスにおける人の認知機能モデルに基づく QoE 向上のためのレート制御手法の提案と評価,” 電子情報通信学会技術研究報告 IN2018-143 (2019).
- [2] H. Nam, K. Kim, and H. Schulzrinne, “QoE matters more than QoS: Why people stop watching cat videos,” in Proceedings of IEEE INFOCOM, pp. 1–9 (2016).
- [3] S. Bitzer, J. Bruineberg, and S. J. Kiebel, “A bayesian attractor model for perceptual decision making,” PLoS computational biology, vol. 11, p. e1004442 (2015).