

E-Learning におけるストリーミング動画教材の視聴状況の分析

片平 昌幸*1

Email: katahira@med.akita-u.ac.jp

*1: 秋田大学大学院医学系研究科医学専攻医療情報学講座

◎Key Words E-Learning システム, 動画配信, ログ解析

1. はじめに

秋田大学大学院医学系研究科では、秋田大学附属病院および学外の病院等に勤務する社会人大学院生に対して、E-Learning システム WebClass を用いた動画配信による講義科目を開設している。

昨年の 2016PC カンファレンスにて、E-Learning システムと Open Source Software (OSS)によるストリーミング動画配信サーバを連携させる教育環境についての実践報告⁽¹⁾を行った。本報告は、そのストリーミング動画配信サーバのアクセスログを解析することにより、動画配信による E-Learning コースを受講した大学院生が、どのような環境から、どのような時間帯に接続したのか、また、動画先頭から最後まで通して視聴したか、あるいは途中で脱落したかなどといった、視聴動向の分析を試みたものである。

2. 医学系研究科における遠隔講義の概要

2.1 E-Learning による遠隔講義

現在、医学系研究科博士課程の必修科目として、遠隔講義を開設している科目の一覧を表 1 に示す。

収録された講義内容の動画データは配信サーバにアップロードされ、全学で共同利用されている E-Learning システムである WebClass 上に開設された各科目のコースの各回の講義のユニットから閲覧できるようになっている。遠隔講義視聴後、それぞれ設定されたレポート課題等を提出することによって各回の講義の受講が完了する。

表 1 遠隔講義開設科目一覧

講義科目名	講義回数
生命科学概論	全 14 回
臨床医学概論	全 16 回
最新医学研究	全 16 回

2.2 動画配信システム

ストリーミング動画配信システムとして、昨年度より HTTP Live Streaming (HLS) プロトコル⁽²⁾を利用した配信サーバを導入した⁽¹⁾。HLS 配信サーバの基本的な構成図を図 1 に示す。HLS による動画配信は既存の各種の技術要素の組み合わせで実現されている。具体的には、H.264/AAC コーデックの MPEG-2 TS 形式の動画ファイルを 10 秒毎に分割したものを、m3u8 形式のプレイリストファイルに記述された再生順に従い、HTTP/HTTPS プロトコルにより配信サーバからダウンロードしつつ再生する。いずれも標準的な規格を用いているので、HLS サーバの構築には専用の有償のものは必要としない。そのため、本システ

ムでは動画形式の変換・分割から配信用 Web サーバまで、全て OSS を用いている⁽¹⁾。

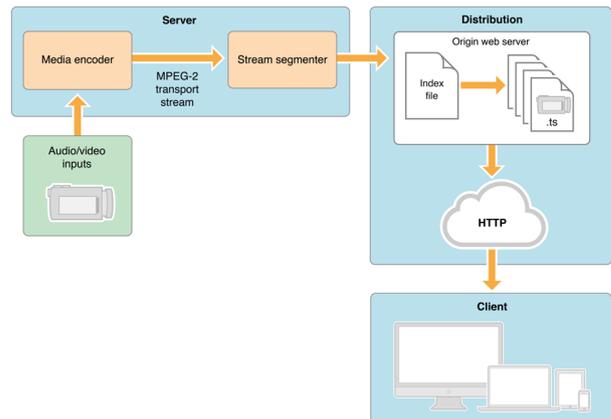


図 1 HLS 配信サーバの基本的な構成⁽²⁾

HLS では、配信サーバとクライアント間の通信には前述の通り通常の Web ページと同じプロトコルを用いているので、ファイアウォールや NAT ルータなどの通信環境の影響を受けにくい。そのため、多種多様なクライアントからの視聴が予想される遠隔講義配信においては視聴トラブルの減少にも役立っている。実際に、以前に利用していた QuickTime Streaming Server (QTSS)配信では、プラグインなどの問題で学生からの再生不可の問い合わせが頻発していたが、HLS に移行後はほとんど無くなった。

HLS による動画配信は HTML5 の video タグに対応していればクライアントの種類を問わず再生可能であるが、Windows PC では Windows10 の Edge ブラウザ以外では標準対応していないという問題がある。これに対応するため、HLS 未対応ブラウザでも HLS 形式動画を再生可能にするための OSS 動画プレーヤとして Video.js⁽³⁾を使用している。

2.3 WebClass 教材への埋め込み

```
<!DOCTYPE html>
<html lang="ja">
  <head>
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width,initial-scale=1,minimum-scale=1,maximum-scale=1,user-scalable=no">
    <title>HTTP Live Streaming Test</title>
  </head>
  <body>
    <div>
      <video src="https://[redacted]/media/hls/sei2015/sei01/sei01.m3u8"
        preload="none" onclick="this.play()" controls />
    </div>
  </body>
</html>
```

図 2 video タグを用いた動画埋め込み例

WebClass 教材における HTML 編集機能を用いて必要な

タグ記述を行うことにより、HLS 配信サーバによる動画配信を教材内に埋め込むことができる。

図2に、HLS 対応ブラウザのための Video タグを用いた動画埋め込み例を示す。

3. HLS 配信サーバアクセスログ解析

3.1 解析前の下処理 (匿名化等)

Web サーバのアクセスログには、接続者のアクセス元の IP アドレスや host 名、使用している機器などの情報が得られる User Agent (UA) 等の、取り扱いに注意を要する情報が含まれている。本報告で対象としている、WebClass サーバと HLS 配信サーバは独立して存在しているため、それぞれの接続ログを突き合わせて確認しない限り、今回の解析で利用した HLS 配信サーバのログのみでは個人の特定は不可能ではあるが、念のため、接続元 IP/host 名および UA 項目等について、解析前の下処理の際に匿名化処理を行った。具体的な処理内容を以下に述べる。

はじめに、Apache Web サーバの Combined フォーマットである HLS 配信サーバのアクセスログから、接続元、UA、リファラの各フィールドのみを個別に取り出して、重複を取り除きユニークなリストを得る。ここからは元のアクセスログとは完全に切り離して処理を行う。

接続元リストに対しては、host 名情報や IP アドレスから接続元のカテゴリ分け (学内、モバイルアクセス、一般プロバイダ、公衆 Wi-Fi) した後、各区分毎に ID 番号を付与して整理する。

同様に UA 情報に関しても、公開されている UA からの各種の OS、ブラウザ判別法に基づいてカテゴリ分けを行い、OS 名と使用ブラウザのみの情報に整理する。

リファラフィールドからは各動画ファイルにどのページからアクセスしたかが判定できる。今回はあらかじめ実際にアクセスして調査しておいたリファラ情報と突き合わせて対応づけることにより、どの教材のどのページからのアクセスか (具体的には、視聴に Video タグか Video.js を使用したか) の判定のみに利用した。

実際の解析に使用した HLS 配信サーバのアクセスログは、オリジナルのログから、講義配信期間である 2016 年 4~12 月の期間における、調査対象の講義の動画データ及び m3u8 プレイリストファイルへのアクセスのみを抽出し、ログ内のアクセス元・UA・リファラの各フィールドは前述の匿名化作業で整理された各リストの項目に置換し、さらにファイルパス名から講義名と、分割された動画ファイルシーケンス番号を抽出したものである。

3.2 アクセスログ解析

前節の下処理にて匿名化したアクセスログを、アクセス元 ID→アクセス日時→講義名の優先順位でキーとしてソートする。このソート処理により、あるアクセス元からの一つの講義に対する一連の視聴動向が、一続きの動画シーケンス番号の列として得られる。このとき、もし、NAT ルータ経由のアクセスなどで、配信サーバからは一つのアクセス元に見えるところから、同時に複数の学生が同一の講義を受講した場合にはシーケンス番号の混乱が発生する可能性があるが、今回の解析ではこのような事象の発生頻度は少ないものとして対応しなかった。

ここで得られた一連の動画シーケンス番号列を、以下

で述べるような方法でさらに精査し、視聴動向の分析を行った。まず、番号列の先頭からは、視聴開始位置が得られる。また、番号列中の最大値からは、動画中で視聴した最後の位置が得られる。番号列に抜けがある (差分が+1より大きい) 場合は、動画途中でスキップしたことが分かる。逆に、時間軸に対して番号列の逆転 (差分がマイナス) が発生している場合は、巻き戻し視聴と判別される。さらに、分割された全ての動画ファイルへのアクセスが記録されている場合は、たとえ途中でスキップや巻き戻しがあつたとしても、動画の先頭から最後まで一度は視聴したことが分かる。これらはいくまでも動画視聴クライアントからのサーバへのアクセスにより判断されるもので、受講学生が集中して視聴していたか否かまでは分からないが、少なくともある程度の視聴動向を評価することは可能であると考えられる。

次節以降のログ解析結果では、視聴動向を表すキーワードとして表2に示すものを用いることとする。

表2 視聴動向キーワード一覧

キーワード	視聴動向
COMPLETED	全ての分割ファイルアクセス有り
FINISHED	98%以上アクセス有り
SKIPPED	最後まで見たがアクセス 98%未満
ABORTED	最後まで見ずアクセス 98%未満
FASTABORT	20%未満しか視聴していない

それぞれ、COMPLETED は完全視聴、FINISHED はほぼ (98%) 視聴済み、SKIPPED は途中飛ばしながらではあるが最後まで視聴、ABORTED は最後まで見ずに途中で視聴を中止した、FASTABORT は早期 (20%未満) の視聴中止に対応している。

4. 解析結果と考察

4.1 視聴環境の概要

ここでは、UA フィールドなどから推測される、動画視聴環境の概要について述べる。

リファラフィールドの解析から、視聴に使用したプレーヤを調べることができるが、今回の分析ではほぼ 100% が Video.js による視聴であった。これは、WebClass 上で作成した教材が「Video.js を使用したページ→前ページが視聴不可の場合は Video タグ使用ページを試みる」という順になっていたのであるが、どの視聴環境でも最初の Video.js 利用のページのほうで問題なく視聴できていたものと考えられる。

図3に、UA から推測した視聴に使用した端末の OS の統計を示す。Windows 系 OS が一番多く 70%あまりを占め、続いて MacOSX (20%弱)、iOS、Android の順となった。さすがに、サポート終了済みの Windows Vista 以前によるアクセスは見られなかった。

図4に、同じく UA から推測した視聴に使用したブラウザの統計、および、図5に、OS 毎の使用ブラウザの内訳を示す。Windows、Macintosh、Android のいずれでも利用可能な Chrome の利用がもっとも多く、次に Apple 社製品限定の Safari、Windows7/8.1 環境で過半数を占める Internet Explorer の順となっている。Windows10 では半数ほどが Chrome であるが、IE によるアクセスはなくその代わりに

Edge が 3 割程度を占めているのが特徴的である。そのほかの割合としては少ないアクセスはタブレット系の動画プレーヤによるものと考えられる。

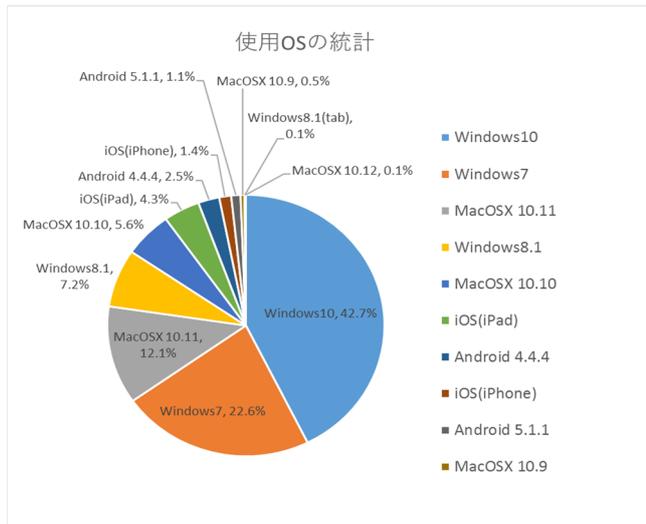


図3 視聴に使用した端末のOSの統計

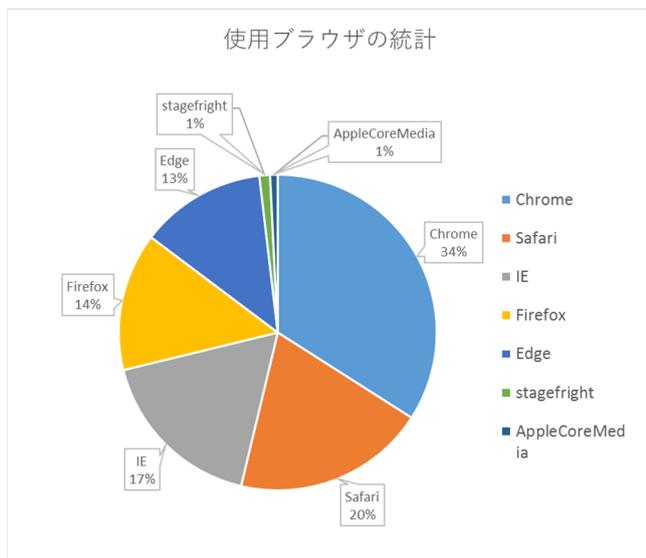


図4 視聴に使用したブラウザの統計

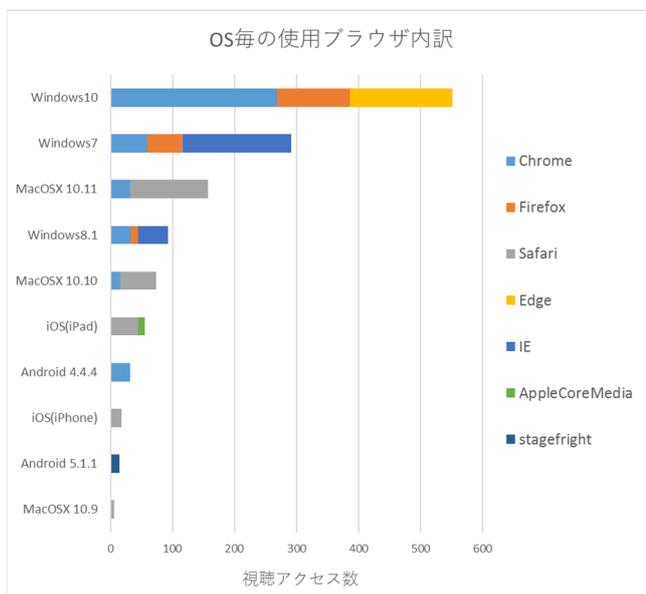


図5 OS毎の使用ブラウザの内訳

4.2 視聴動向の分析

動画ファイルアクセス状況に基づく、視聴完了状況の分析結果を図6に示す。70%強が完全視聴であり、受講生はかなり真面目に視聴していることがうかがわれる。

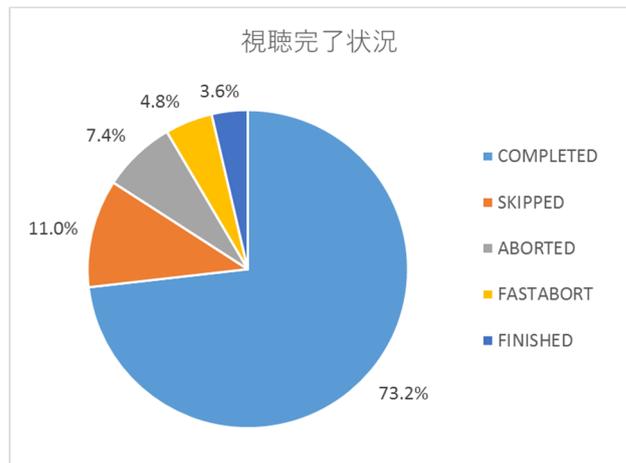


図6 視聴完了状況

図7に、匿名化処理後の接続元種別毎に、それぞれの接続元からの視聴総数に対する視聴完了状況を分析した図を示す。秋田大学学内LANからのアクセスがもっとも完全視聴の割合が高く、一般プロバイダからモバイルアクセスの順に途中中断や早期中断が増加するという、接続元から推測される視聴時の周囲環境の影響との関連性が強く示唆される結果が確認できた。なお公衆Wi-Fiからのアクセスは今回サンプル数が非常に少なかったため、視聴完了状況との関連性についてはさらなる調査分析が必要である。

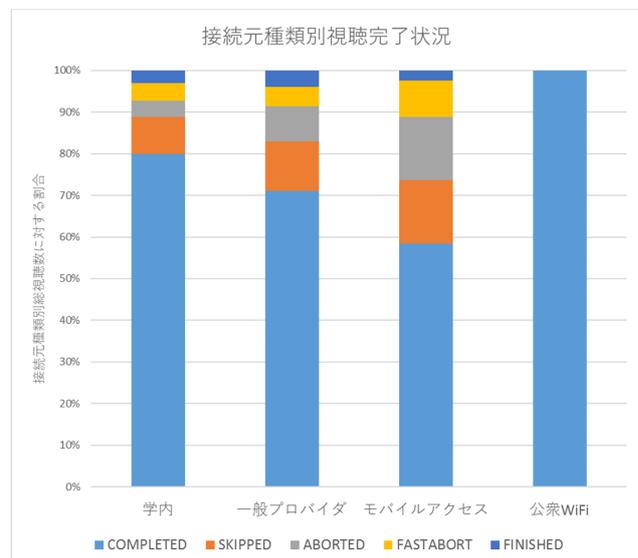


図7 接続元種別別視聴完了状況

4.3 月別・時間帯別視聴状況

図8に、接続元種類毎の月別視聴アクセス数の状況を示す。本報告の分析対象とした遠隔講義配信期間は2016年4月から12月末までであるが、2016年度にあらたに配信開始した講義がいくつかあり、これらは配信期間内に順次撮影と公開を行っている。しかしながらもっとも大きな影響は、配信終了となる年末間際の駆け込み視聴で

あることがこのグラフから容易に推測できる。

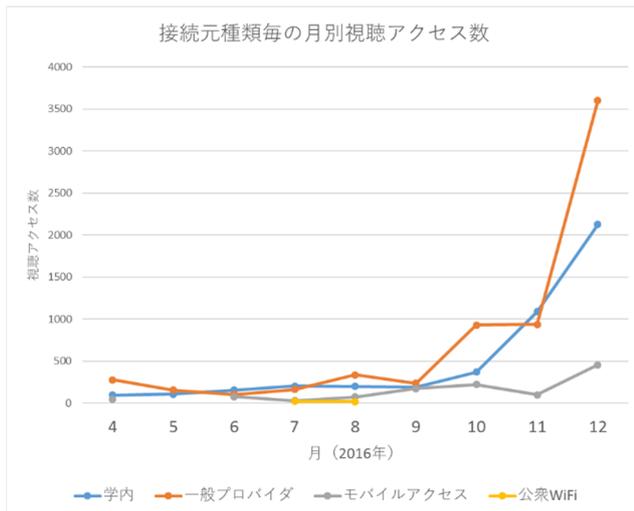


図8 接続元種類毎の月別視聴アクセス数

図9に、接続元種類毎の時間帯別視聴アクセス数の状況を示す。アクセス元の分類における一般プロバイダには個人契約と勤務先の契約がともに含まれるが、プロバイダのhost名からはこれらは分類不可能であるので合計して分類している。図9のアクセス状況から、一般プロバイダからのアクセスは、おそらく日中の勤務先からのアクセスと夜間の自宅からのアクセスとが重なっていることが推測できる。また、モバイルアクセスも夕方以降増加していることが分かる。なお、解析に使用した匿名化加工済みログデータを実際に目視確認したところ、時間帯別における21時以降の急増分には、12月末日のぎりぎりの駆け込み視聴の影響もあることがわかった。

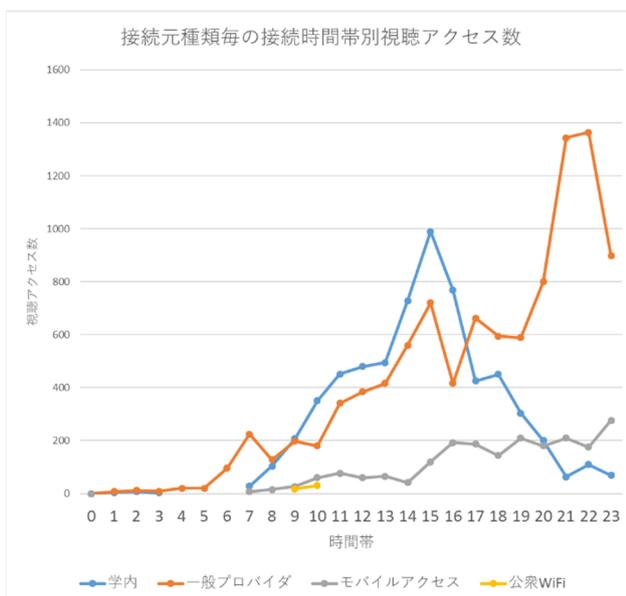


図9 接続元種類毎の時間帯別視聴アクセス数

4.4 配信動画の長さとの視聴動向

動画配信をもちいた遠隔講義では、配信される動画の時間が長すぎると視聴者が途中で飽きてしまい、視聴を中断してしまう懸念がある。そのため、配信する講義動画の長さ(HLS配信では10秒毎に分割された動画ファイルの総数から容易に求めることが可能)と視聴動向との関

連について分析を試みた。図10に、視聴動向(完全視聴と飛ばし見あり)と教材動画長との関連性について描いた散布図を示す。教材動画が扱っている講義内容に対する関心度の影響も一部あるとは考えられるが、動画長が長くなるに従い、明らかに完全視聴の割合は下がり、飛ばし見視聴の割合が増加していることが確認された。

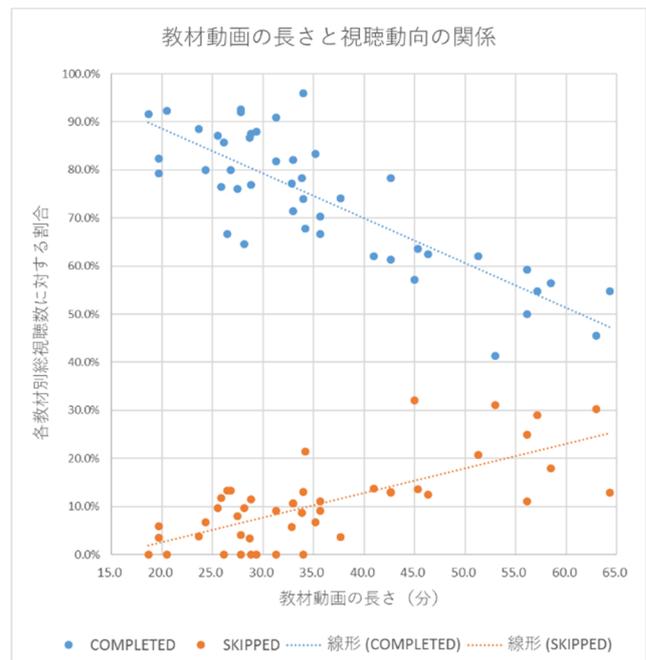


図10 配信教材動画長と視聴動向の関係

5. おわりに

本報告では、HLS 型式による動画配信サーバと E-Learning システムを連携して開講した遠隔講義について、配信サーバのログを解析することにより、受講者の視聴環境や視聴動向についての分析を試みた結果について述べた。OS やブラウザなどの実視聴環境の分析から、配信システムの構築の際に対応が必要な視聴環境を改めて確認することができた。今後は、今回得られた視聴環境や動画長などと受講者の視聴動向との関連の分析を元にして、よりよい遠隔講義環境の構築を目指したいと考えている。

参考文献

- (1) 片平昌幸: “EラーニングシステムとOSSを用いたストリーミング動画配信サーバの連携の実践報告”, 2016PCカンファレンス論文集, pp.175-176 (2016).
- (2) Apple Inc., HTTP ライブストリーミングの概要, <https://developer.apple.com/jp/documentation/StreamingMediaGuide.pdf>
- (3) <http://videojs.com/>