

ライブ動画配信におけるアクセス傾向

神屋 郁子[†] 下川 俊彦[‡]

[†]九州産業大学大学院情報科学研究科 〒813-0011 福岡市東区松香台 2-3-1

[‡]九州産業大学情報科学部 〒813-0011 福岡市東区松香台 2-3-1

E-mail: [†] kamiya@nw.is.kyusan-u.ac.jp, [‡] toshi@is.kyusan-u.ac.jp

あらまし 多数のクライアントに安定した動画配信を行うため、ライブ動画配信のログを利用し、アクセス傾向の調査を行った。本論文では、ログの記録率、視聴アクセスの割合、視聴用ソフトウェアについて調査した結果を報告する。

キーワード ストリーミング, アクセスログ, メディアプレイヤー

Analysis of Live Streaming Server Access Log

Yuko KAMIYA[†] Toshihiko SHIMOKAWA[‡]

[†] Graduate School of Information Science, Kyushu Sangyo University 2-3-1 Matsukadai, Higashi-ku, Fukuoka, 813-0011 Japan

[‡] Faculty of Information Science, Kyushu Sangyo University 2-3-1 Matsukadai, Higashi-ku, Fukuoka, 813-0011 Japan

E-mail: [†] kamiya@nw.is.kyusan-u.ac.jp, [‡] toshi@is.kyusan-u.ac.jp

Abstract In order to deliver stable live streaming for many clients, we analyze access tendency using live streaming server access log. In this paper, we report three things; the rate that log was recorded, the rate of watching access, and software for watching.

Keyword Streaming, Access Log, Media Player

1. はじめに

近年、日本では、いわゆるブロードバンドネットワークが急速に普及した。この結果、大容量のデータが容易に転送できるようになった。ブロードバンドネットワークの普及で広く用いられるようになった応用の1つに動画配信（ストリーミング）がある。本研究が対象とする動画配信は、全てストリーミングによるものである。動画配信の中には、多数のアクセスが見込まれるコンテンツも多く、多数のクライアントに安定した映像を配信する必要がある。動画配信の品質を計る手法の1つに、サーバのログを分析する手法がある[1]。しかし、実際の動画配信のログを解析した結果、十分な情報が記録されていない場合が少なくないことがわかった。

そこで、本研究では、実際の動画配信におけるアクセスログを調査した。まず、ステータスコードに注目すると、正常終了を意味する 200 以外が記録されているアクセスが非常に多いことがわかった。これは、正常に視聴ができていない可能性を示すと考えた。そこで、これらのアクセスについて調査を行った。この結果、正常終了はしていないものの、視聴はできている

場合が多いことがわかった。

また、この調査の中で、サードパーティ製再生アプリケーションを用いている利用者が居ることが多いことがわかった。この場合、正常終了しないことがわかった。正常終了しないので、前述のようにアクセスログに、十分な情報が残らない。そこで、実際の動画配信のログを解析し、どの程度の利用者がサードパーティ製アプリケーションを利用して視聴しているかの調査を行った。

本論文では、これらの調査結果について報告する。

2. 動画配信システム

現在普及している動画配信システムは、開発元が配信側アプリケーションから再生側アプリケーションまでを一括して提供している場合が多い。これは歴史的に、動画配信システムが動画の codec や配信プロトコルに独自のものを利用することが多かったことに起因する。この結果、多くの利用者は、開発元が提供する再生アプリケーションを利用して視聴する。

しかし、開発元が提供する再生用アプリケーションにない機能を追加した再生用アプリケーションをサー

ドパーティが提供する場合や、そもそも再生用アプリケーションが提供されていないプラットフォーム用に、サードパーティが再生用アプリケーションを提供する場合があります。4.2 で述べるように、このようなサードパーティ製再生用アプリケーションは、アクセスログに十分な情報を残さないことがあることがわかった。おそらく、必要な情報が得られないまま実装されたため、サーバ・クライアント間での十分な情報の通知が実装されていないためではないかと考えている。

本論文では、以下、再生用アプリケーションのことをメディアプレイヤーと呼ぶ。

2.1. アクセスログ

動画配信システムは、それぞれ独自のアクセスログの書式を持つ。本研究では、Windows Media サービス 4.1 のアクセスログを解析対象としたので、その書式について説明する。Windows Media サービスの最新のバージョンは 9 である。本研究で 4.1 を対象とした理由は、3.2 で述べるように、本研究の解析対象とした配信システムが Windows Media サービス 4.1 を用いて構築されたためである。

2.1.1. フィールドについて

Windows Media Server のアクセスログには全部で 43 のフィールドが記録される[2]。本研究では、この中から以下の 4 つのフィールドに着目して解析を行った。

- ・ **c-status** : アクセスの終了状態を表すコード。RFC2326[3] に基づいている。詳細は 2.2.2 で述べる。
- ・ **cs(User-Agent)** : アクセスに使用された Web ブラウザまたはメディアプレイヤーの種類
- ・ **c-playerid** : メディアプレイヤーのグローバル意識別子(GUID)
- ・ **c-hostexe** : クライアントマシン上で実行された、Web ブラウザまたはメディアプレイヤーの実行ファイル名
- ・ **x-duration** : ファイルの再生秒数

2.1.2. ステータスコード

表 1. ステータスコードの種類

ステータスコード	意味
1xx	情報提供
2xx	成功
3xx	転送
4xx	クライアントエラー
5xx	サーバーエラー

ステータスコードとは、各アクセスの終了状態を表す数値のことである。動画配信サーバのログに記録さ

れているステータスコードは、RFC2326 で定義されている RTSP のステータスコードと同様のものとされている。ステータスコードは 3 桁の数値で表される。最上位の数値で表 1 の 5 種類に分類される。

3. 動画配信システムのアクセス解析

本研究では、実際の動画配信の際のアクセスログを対象に、解析を行った。

3.1. 解析対象

本研究では、LIVE! UNIVERSE プロジェクト[4]が実施した 2 回の日食中継の動画配信のログの解析を行った。また、様々なメディアプレイヤーがどのようなアクセスログを残すかを調べるために、実験用の動画配信システムを構築し、そのログの解析を行った。

3.2. LIVE! UNIVERSE

LIVE!UNIVERSE は、天文および宇宙科学などに関する様々な現象やイベントをネットワークを通じて広く社会に紹介し、社会に貢献することを目的としている団体である。本研究では、2005 年 10 月 3 日に行われた LIVE! ECLIPSE 2005 Annular 金環日食 (LE2005a) と、2006 年 3 月 29 日に行われた LIVE!ECLIPSE 2006 皆既日食 (LE2006) の動画配信サーバのアクセスログについて解析を行った。また、比較のため、LIVE!ECLIPSE 2006 の Web サーバのログも解析した。

3.2.1. LIVE!ECLIPSE 2005 Annular

LIVE!ECLIPSE 2005 Annular は、2005 年 10 月 3 日(月) 16:30~18:30 にマドリッドで観測された日食をライブ配信した[5][6]。観測地点から日本へ伝送された映像を元に番組を制作した。制作された映像は Windows Media サービス 4.1 用にエンコードし、StarBED に送信した。StarBED とは、独立行政法人情報通信研究機構北陸リサーチセンター(旧北陸 IT 研究開発支援センター)である。512 台の PC クラスタとそれを接続するスイッチで構成された大規模なネットワーク実験環境である。StarBED の PC には、Windows 2000 Server がインストールされていた。これが、本研究が Windows Media サービス 4.1 を対象とした理由である。

視聴方法は、Web ブラウザから、動画配信へのリンクをクリックすると、メディアプレイヤーが立ち上がる形式が採用された。

3.2.2. LIVE!ECLIPSE 2006

LIVE!ECLIPSE 2006 は、2006 年 3 月 29 日(水) 18:00~20:30 にリビア・トルコ・エジプトの 3 箇所で観測された日食をライブ配信した。観測地で配信用にエンコードし、その映像を LE2005a 同様 StarBED から配信した。

リビアの映像は Quick Time 形式で配信した。トルコ・エジプトの映像は Windows Media サービス 4.1 形

式で配信した。本研究では、トルコ・エジプトの映像の配信のログのみを解析対象とした。

視聴方法は、Web ブラウザから、動画配信のページにアクセスをすると、Web ページに埋め込まれたメディアプレイヤーが再生を行う方式が採用された。また、1つの画面にトルコ・エジプトの2つの映像が埋め込まれた

3.3. 実験用動画配信

様々なメディアプレイヤーがどのようなアクセスログを残すかを検証するために実験用の動画配信を行った。実験環境は、LE2005a、LE2006 と同様にするため、Windows Media サービス 4.1 を利用した。

4. アクセス解析結果

1. で述べたように、本研究では、LE2005a と LE2006 の動画配信サーバのログの解析を行った。

4.1. ステータスコードごとのログ

まず、各ログについて、ステータスコードに注目して解析を行った。

4.1.1. Web サーバと動画配信サーバの比較

解析を行った動画配信サーバの2種類のログには、主に3種類のステータスコードが記録されていた。それぞれのステータスコードと意味を表2に示す。

表2：ステータスコードと意味

ステータスコード	意味
200	OK
400	Bad Request
408	Request Time-out

まず、LE2006 の Web サーバと動画配信サーバにおけるステータスコードの割合を図1に示す。

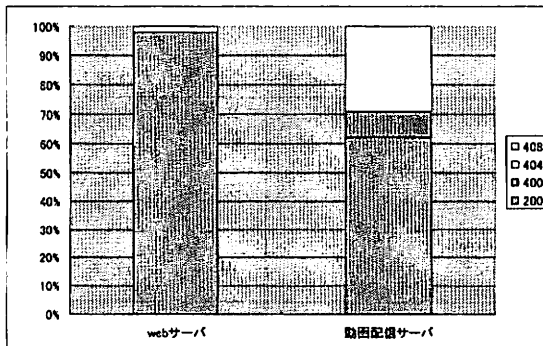


図1：LE2006 動画配信サーバ

図1から、Web サーバのステータスコードはほとん

どが200であるのに、動画配信サーバのステータスコードは、200であるものが半数程度しかなく、400・408の割合が高いということがわかる。

4.1.2. ログの記録状況

ステータスコードごとのログの残り方について調査した。[1]では、クライアントの品質推定を行う際には、回復パケット数、損失パケット数の記録を観測することが重要であると述べられている。そこで、本研究では、回復パケット数、損失パケット数の記録率について注目した。しかし、[1]では、どのフィールドを対象にしたかが具体的に述べられていない。そこで、回復パケット数、損失パケット数が記録されている全てのフィールドを対象とした。具体的には、回復パケット数については、c-pkts-recovered-ECC、c-pkts-recovered-resentの2つのフィールド、損失パケット数については、c-pkts-lost-client、c-pkts-lost-net、c-pkts-lost-cont-netの3つのフィールドを対象として調査した。

表3、表4では、ステータスコードが200・400・408のときに、5つのフィールドごとに記録が残っているアクセスの数を調査した。また全体のアクセス数を元にして、記録されているアクセスの割合を調査した。

表3：LE2005aにおけるログの記録状況

	200	400	408	記録率
recovered-ECC	82911	11	0	43%
recovered-resent	82911	11	0	43%
lost-client	82911	44	21458	54%
lost-net	82911	11	0	43%
lost-cont-net	82911	44	21458	54%

表4：LE2006におけるログの記録状況

	200	400	408	記録率
recovered-ECC	413466	328	0	61%
recovered-resent	413466	328	0	61%
lost-client	413466	36756	43004	73%
lost-net	413466	328	0	61%
lost-cont-net	413466	36756	43004	73%

その結果、LE2005a・LE2006共にこれらのフィールドの記録が残っていないアクセスログが残っていること、さらにその割合が少なくないことがわかった。ま

た、ステータスコードごとに分類すると、ステータスコードが 400 や 408 のときのほうが、200 のときに比べて、これらのフィールドが記録されないアクセスログが多いことがわかった。

4.2. ステータスコードごとの視聴アクセスの割合

4.1.1 で述べたとおり、動画配信サーバには正常でないステータスコードを示すものが多い。そこで、各アクセスごとの視聴時間を調査した。そのためにログから、x-duration のフィールドに注目した。このフィールドはクライアントで動画データが再生された時間である。本研究では、x-duration が 0 でないものを視聴アクセスとする。ステータスコードごとの視聴アクセスの割合を調査した。その結果を表 5、表 6 に示す。

表 5 : LE2005a におけるステータスコードごとの視聴アクセスの割合

ステータスコード	視聴アクセス数	総アクセス数	視聴アクセスの割合
200	101019	82918	55%
400	50	30983	1%
408	57843	76674	76%

表 6 : LE2006 におけるステータスコードごとの視聴アクセスの割合

ステータスコード	視聴アクセス数	総アクセス数	視聴アクセスの割合
200	45780	198114	24%
400	530	63254	1%
408	100989	413466	25%

表 5、表 6 に示すように、ステータスコードが 408 のものにも、視聴アクセスがあることがわかった。しかし、表からわかるように、視聴アクセスの割合は高くない。特に、LE2006 の割合は高くない。しかし、これは問題ではないと考えているなぜなら、ステータスコードが 200 の場合の視聴アクセスの割合と同様だからである。よって、この問題の原因はクライアント側でなく、サーバ側、あるいは配信されるコンテンツの問題ではないかと考える。

4.3. 視聴用ソフトウェア

4.1 でステータスコードが 200 でないアクセスが多いということがわかった。マイクロソフト社のドキュメントによると、以下のような場合にログに 408 が記録されると述べられている [7]。

- Windows Load Balance Service (WLBS) を利用しており、WLBS セットアップで類似性を無効にしているか [None] に設定している
- AOL ブラウザにプレーヤーが埋め込まれて

いる

- Macintosh プレーヤーが接続されている
- 再生中にプレーヤーを閉じた
- プレーヤーがファイアウォールの内側にある
- Web ページに埋め込まれたプレーヤーを使用している
- サーバーのユニキャスト ステーション ストリームが停止している
- ロードバランサーのヘルスチェックで利用されているような SYN アクセスがある

しかし、我々は、これら以外にサードパーティ製のメディアプレーヤーを使っていることがその一因ではないかと考えた。そこで、サードパーティ製のメディアプレーヤーがどのようなログを残すかを調査し、実際の配信ログと比較を行った。

4.3.1. メディアプレーヤーが残すログ

サードパーティ製メディアプレーヤーとして以下のものを対象とした。

- Net Transport
- SDP
- GetASFStream
- WMRecorder

NetTransport・SDP・GetASFStream・WMPlayer は、いわゆるダウンロードツールと呼ばれるものである。NetTransport・SDP・GetASFStream は国内で広く知られているツールである。また、WMRecorder は海外で利用されているダウンロードツールである。

正常なアクセスとの比較を行うため、Windows Media Player でのアクセスも行った。

また、様々な条件でのアクセスログを調査するために、各メディアプレーヤー正常終了させる場合と正常終了させない場合について調査した。正常終了させない方法としては、LAN ケーブルを抜いて強制的にネットワークを切断させることで終了させる方法を用いた。

この結果を表 7 に示す。実験により、各ダウンロードツールは同じ挙動を示したのでまとめて記載した。

表 7 : メディアプレーヤーごとのログの残り方

視聴用ソフトウェア	正常終了時ステータスコード	異常終了時ステータスコード	ステータスコード以外のフィールド
Windows Media Player	200	408	違う
ダウンロードツール	408	408	同じ

Windows Media Player を利用したアクセスでは、正常終了時にはステータスコードが 200 になる。異常終了時にはステータスコードが 408 になる。また、異常終了時には例えば c-id フィールドが記録されないなど、ステータスコード以外のフィールドも正常終了時とは違うことがわかった。

ダウンロードツールを用いたアクセスでは、正常終了時も異常終了時にもステータスコード 408 が記録される。また、ステータスコード以外のフィールドも正常終了時でも異常終了時でも同じである。

この結果から、ステータスコード 200 になるものは Windows Media Player で正常に終了したとき、ステータスコードが 408 になるのは、Windows Media Player で異常終了したとき、あるいは、ダウンロードツールを利用したときであるということが推測できる。

次に、それぞれのメディアプレイヤーが残すログの特徴を表 8 に述べる。表中 WMP とは、Windows Media Player のことである。

表 8：メディアプレイヤーごとのログの特徴

視聴方法	cs(User-Agent)	c-id
WMP での視聴	○	○
WMP を異常終了させる	NSPlayer/c-id	○
WMP を異常終了させる	NSPlayer	○
SDP	NSPlayer	×
GetASFStream WMRecorder NetTransport	×	○

Windows Media Player で視聴した場合、cs(User-Agent)に記録されるデータは、配信ページによって 2 種類に分類される。メディアプレイヤー埋め込み型とメディアプレイヤー独立型である。メディアプレイヤー埋め込み型とは、web ブラウザにメディアプレイヤーが埋め込まれている場合である。メディアプレイヤー独立型とは、メディアプレイヤーが立ち上がり再生が行われる場合である。3.2.1、3.2.2 で述べたように、LE2005a では埋め込み型、LE2006 では埋め込み型でない配信ページが作成された。埋め込み型の場合、cs(User-Agent)には、配信ページへのアクセスを行ったブラウザの User-Agent が記録される。メディアプレイヤー独立型の場合、再生したメディアプレイヤーの User-Agent が記録される。また、c-id には、利用した

メディアプレイヤーの id が記録される。しかし、Windows Media Player で異常終了させると、cs(User-Agent)には NSPlayer と c-id が記録され、c-id には記録があった。また、SDP では、cs(User-Agent)に NSPlayer が記録されたが、c-id には何も記録されなかった。また、GetASFStream、WMRecorder、NetTransport を利用すると、cs(User-Agent)には何も記録されなかったが、c-id には記録があった。

以上の点を元に、ステータスコード 408 のログを分類した cs(User-Agent)に NSPlayer/c-id が記録された場合と cs(User-Agent)に NSPlayer と記録されており、c-id に記録があるものは、表 8 より Windows Media Player を異常終了させた場合と判断した。cs(User-Agent) に記録があり、c-id にも記録があるログが残っていた。これは、表 8 には現れない場合である。しかし、status コード 200 の場合のログとの照合から、この場合も Windows Media Player の異常終了と判断した。

cs(User-Agent)か、c-id のいずれかに c-id が記録された cs(User-Agent) が NSPlayer と cs(User-Agent) に NSPlayer と記録されているものは SDP と判断した。cs(User-Agent) が記録されておらず、c-id が記録されているものは、GetASFStream、WMRecorder、NetTransport のいずれかだと判断した。最後に cs(User-Agent)に記録が無く、c-id も記録されていないものは、不明なアクセスとした。これは、このパターンのログは、ほとんどのフィールドに記録が残っておらず推測の材料がないためである。

考察したログの分類方法を元に、LE2005a、LE2006 のステータスコード 408 のログを分類した。実数と割合を表 9、表 10 に示す。

表 9：LE2005a におけるステータスコード 408 のツールのアクセス実数と割合

視聴方法	実数	割合
WMP で視聴中 LAN を抜く	73481	97%
WMP を利用したアクセスでは？	0	0%
SDP	916	1%
GetASFStream WMRecorder NetTransport	1813	2%
不明なアクセス	235	0%

表 10：LE2006 におけるステータスコード 408 のツールのアクセス実数と割合

視聴方法	実数	割合
WMP で視聴中 LAN を抜く	191332	96%
WMP を利用したアクセスでは？	25	0%
SDP	1224	1%
GetASFStream WMRecorder NetTransport	5248	3%
不明なアクセス	285	0%

この結果から、利用者のほとんどが Windows Media Player を利用しているということがわかる。また、サードパーティ製再生アプリケーションを用いている利用者も 4%ほどいるのではないかと推測できる。

5. おわりに

本研究では、ライブ動画配信におけるアクセスログについて 3 つの調査を行った。1 つ目に、ステータスコードに注目し解析を行った。その結果、約半数のアクセスログにおいて、ログに記録されるべきフィールドのうちのいくつかが未記録となっていることがわかった。これは、サーバの仕様によるところでもある。しかし、これらのフィールドはクライアント品質推定に有効という研究結果も出ており、留意が必要と考える。

次に、正常終了を意味しないステータスコードが多いことに注目し、ステータスコードごとの視聴時間の調査を行った。この結果、視聴時間はステータスコードには依存しないことがわかった。

最後に、ログの特徴を元に利用者が使っているメディアプレイヤーの推測を試みた。この結果、約 4% 程度の利用者が、いわゆるダウンロードツールを利用していることが推測できた。ストリーミングの特徴の 1 つに、配信コンテンツが利用者の手元に残らないということがある。しかし、今後ダウンロードの利用が広まるようであれば、DRM などの利用が必須になると考える。

謝辞

本研究の一部は科学研究費補助金、特定領域研究 18049009、および基盤研究 (B) 17300012 の支援を受けている。

NTT スマートコネクトの谷崎文義氏、遠藤論氏、松

下電器産業の東正美氏、ブロードバンドタワーの芝野智誉氏からは、様々な有意義なアドバイスを頂いた。

III の山本文治氏を初めとする Streams-JP メイリングリストの参加者のみなさんには貴重なご意見を頂いた。

文 献

- [1] “クライアント品質推定のための詳細なストリーミングサーバログ分析”, 寺田直美, 河合栄治, 砂原秀樹, 情報処理学会論文誌, 第 47 巻第 7 号, pp.2019-2029, July 2006.
- [2] “[netshow] Windows Media ログ ファイルの参考資料”, Microsoft Corporation, <http://support.microsoft.com/kb/285847/ja>
- [3] “Real Time Streaming Protocol (RTSP)”, H. Schulzrinne, A. Rao, R. Lanphier, RFC2326, April 1998
- [4] LIVE! UNIVERSE, <http://www.live-universe.org/>
- [5] “LE2005a における StarBED を利用した大規模ストリーミングサーバの構築”, 下川 俊彦, 池野桂司, 池上 恵介, 東 正美, 青木 哲郎, 情報処理学会研究報告, 2006-QAI-18, pp. 1 - 6, Jan 2006
- [6] “LE2005a の概要とマドリッドからの高速回線による高画質映像伝送”, 山本文治, 中村一彦, 尾久土正己, 情報処理学会研究報告, 2006-QAI-18, pp. 13 - 18, Jan 2006
- [7] <http://support.microsoft.com/kb/256277/ja>