VOL. 4 NO. 1 DECEMBER 2019

JOURNAL OF DIGITAL VIDEO

SCTE • ISBE STANDARDS

DIGITAL VIDEO

scte.org • isbe.org

SCTE · ISBE

Society of Cable Telecommunications Engineers
International Society of Broadband Experts

JOURNAL OF DIGITAL VIDEO

VOLUME 4, NUMBER 1
December 2019

Society of Cable Telecommunications Engineers, Inc.
International Society of Broadband Experts™
140 Philips Road, Exton, PA 19341-1318

Table of Contents

Editorial Correspondence: If there are errors or omissions to the information provided in this journal, corrections may be sent to our editorial department. Address to: SCTE Journals, SCTE•ISBE, 140 Philips Road, Exton, PA 19341-1318 or email journals@scte.org.

Submissions: If you have ideas or topics for future journal articles, please let us know. Topics must be relevant to our membership and fall under the technologies covered by each respective journal. All submissions will be peer reviewed and published at the discretion of SCTE•ISBE. Electronic submissions are preferred, and should be submitted to SCTE Journals, SCTE•ISBE, 140 Philips Road, Exton, PA 19341-1318 or email journals@scte.org.

Subscriptions: Access to technical journals is a benefit of SCTE•ISBE Membership. Nonmembers can join at www.scte.org/join.

Welcome to the December issue of the *Journal of Digital Video*, a publication of collected papers by the Society of Cable Telecommunications Engineers (SCTE) and its global arm, the International Society of Broadband Experts (ISBE). This edition of the journal focuses on various aspects of the digital video sector.

The first article gives an overview of a new method for effective quantitative audience measurement The most popular audience tracking methods in use were developed in the era of dominant linear video distribution and have limitations in understanding the performance of nonlinear viewing, spread across a panoply of distribution methods and devices. This paper describes a new method, trackable asset cross-platform identification (TAXI), which was developed jointly by the Coalition for Innovative Media Measurement (CIMM) and the Society of Motion Picture and Television Engineers (SMPTE), and defines a standard way of inserting unique identifiers into any different part of audiovisual streams whether content or advertisement.

Video piracy continues to be a significant concern for content creators and distributors, with piracy models, anti-piracy technologies and best practices continuing to evolve as awareness increases. The second paper explains piracy use cases and describes anti-piracy initiatives as a technical process that must be complemented by well-informed business rules and business policies that guide the selection and use of anti-piracy technology.

The third paper describes novel applications of blockchain technology in digital TV advertising and alternate content switching based on the auditing requirements in the recently created standard SCTE 224 2018, "Event Scheduling and Notification Interface (ESNI)." SCTE 224 offers rich capabilities to support a wide variety of alternate content and advertising scenarios and, as programmers introduce new features, content distributors are expected to support them. However, it also is known in the industry that validation of content switching in IP streaming is a formidable challenge and sending out large amounts of customer device data raises privacy concerns as well. The paper describes a blockchain-based solution to address this intractable issue.

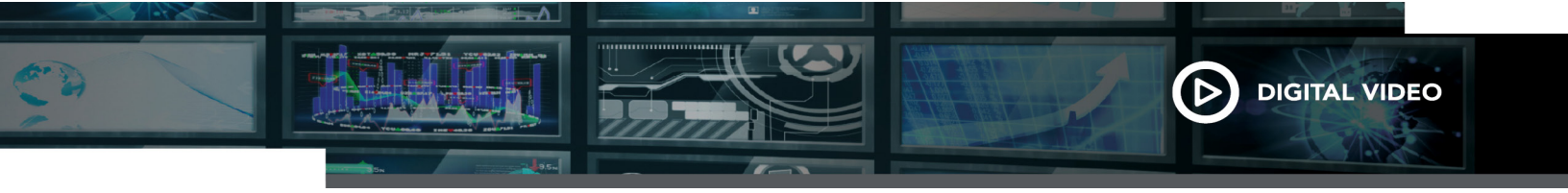
In the final paper, encoding intelligence for live video distribution is explored. Real-world live video distribution systems are faced with the challenge of processing videos of extremely diverse content types and complexity in real time. To avoid severe and unpredictable quality variations across time, video assets, and content types, the authors describe a quality-of-experience (QoE) metric that predicts end viewers' experience when consuming videos streamed to their viewing devices, with consistent QoE predictions across content type, content complexity, codec type, bit rate, video resolution, frame rate and dynamic range.

We thank the individuals who contributed to this issue of the *Journal of Digital Video*, including the authors, peer reviewers, and the SCTE•ISBE publications and marketing staff. We hope you enjoy this issue and that the selected papers spark innovative ideas and further cement essential knowledge in digital video.

In closing, if there is any editorial information or topics that you would like us to consider for the next issue of *SCTE•ISBE Journal of Digital Video*, please refer to the "editorial correspondence" and "submissions" sections at the bottom of the table of contents for instructions.

SCTE•ISBE Journal of Digital Video Editor,

Paul Hearty, Ph.D. (Senior Editor)
Technology Advisors
SCTE Member



A New Method for Effective Quantitative Audience Measurement

Trackable Asset Cross-Platform Identification

A Technical Paper prepared for SCTE•ISBE by

William Kreth, ex-Executive Director – Entertainment Identifier Registry Association (EIDR),
Consultant / Advisory Board Member, Simply.TV
248 Park Pl.
Brooklyn, NY 11238
w.e.kreth@gmail.com
(917) 714-4532

Francois Modarresse, Consultant, Entertainment Identifier Registry Association (EIDR),
SCTE•ISBE Member
C/O MESA 1416 N 13th St.,
Boise, iD 83072
fmod@advemtv.com
(702) 984 6000

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Abstract

The rising consumption of video and the even faster growing number of mediums via which it is delivered, are gradually causing the obsolescence of television audience measurements. Non-linear viewing, non-standard format length, proliferation of highlights, coexistence of single-to-multipoint with multi-point to multipoint distribution are turning the traditional practices into approximations at best.

Neither manual polling, nor voluntary reporting, and not even fixed monitoring facilities will respond to this shift. Solid, novel techniques are required to accurately measure. Embedding tracking information within the content itself has gained popularity to respond to the challenge. However, the use of disparate, uncoordinated method poses a problem of multi-modal detection of content and its relation to the original source.

To address the issue, the Coalition for Innovative Media Measurement (CIMM) and the Society of Motion Picture and Television Engineers (SMPTE) have teamed-up to jointly develop a standard way of inserting unique identifiers into any different part of audiovisual streams whether content or advertisement.

This new method is called TAXI (trackable asset cross-platform identification) Complete.

In this paper, we will describe how this new standard enables new efficiencies by unifying solutions and processes. It saves cost by reducing the need for the endless development of new or dedicated equipment particularly at the endpoint.

We will also show how the recommended methods survive content encoding, transcoding and editing across processing, transmission and playback platforms.

Finally, we will relate how live demonstrations of TAXI Complete made on emerging broadcast networks such as the Advanced Television Systems Committee (ATSC) 3.0 have produced measurable results that address many of the limitations of standard methods, and how additional development on existing distribution networks may further drive adoption and long-term value of the industry. These demonstrations are a clear precursor of how similar techniques could be extended to cable television and should favor audience measurement for linear and non-linear viewership across devices.

1.2. Innovation Background

Video production, consumption and revenues in the media and entertainment (M&E) sectors continue to enjoy a robust 5% annual growth (Forbes 2019). Considering the thousands of sources that already existed in established multichannel video programming distributor (MVPD) networks, and the explosion of niche providers (Forbes), the trends are expected to endure, if not amplify.

Concurrently, the methods to distribute and consume the M&E content are proliferating (PWC). Beyond the profusion of sources to access content, a completely open choice between linear and nonlinear content has become the new normal (Statista).

The M&E industry has traditionally used audience tracking as one of the primary tools to understand their consumers' habits and their own business performance to drive their business decisions. It has also been an essential quantifier to drive or reconcile revenue streams such as advertisements.

The most popular audience tracking methods were developed in the era of dominant linear video distribution. These techniques although still largely in place, show limitations in understanding the performance of nonlinear viewing, spread across a panoply of distribution methods and devices. They need therefore to be augmented, if not replaced, with tools that better capture contemporary consumption behaviors.

2. Historical Methods For Audience Measurement

Over the years, audience measurement was performed in several different ways depending on technologies available then and the specifics of the video service to be tracked.

2.1. Audience Polling

Simply conducted by operators calling audience panels to ask them what they watched.

2.2. Opt-In Household Reporting

A service provider coordinating audience measurement, designed via a voluntary panel of consumers to track their viewing (e.g.- using a paper-based viewing diary and/or a dedicated device to detect what content is/was being watched - such as a "portable people meter"). (Wikipedia)

2.3. Monitoring facilities

A series of receivers and associated monitoring equipment installed in a professional facility. This particular method does not apply to audience measurement per se. It is more dedicated to compliance check. It is however listed here, as modern audience measurements can also be repurposed for compliance.

2.4. STB / DVR Data Mining

Set top box (STB) receivers and digital video recorders (DVR) include capabilities to anonymously (or through a process of de-identification) log and report their users' television programming consumption and can therefore be used for audience measurement.

In the context of this study, these different mechanisms can be qualified by the level of automation that they are based on, the consumer's degree of involvement that they entail, and the amount of dedicated equipment they require.

2.5. Comparative analysis of legacy methods

Table 1 – Legacy Methods

Method	Automation	Consumer involvement	Equipment needed	Remarks
Audience Polling	Fully Manual	High	Standard	Labor intensive and not scalable.
Opt-In Household Reporting	Manual to Partially Automated	Medium to High	Dedicated	Requires considerable logistics of shipping and maintaining viewing diaries and/or detection equipment to a consumer home and collecting / processing disparate responses via analog diaries and device data logs. Does not track out-of-home and non-TV viewing.
Monitoring facilities	Automated	None	Dedicated	Compliance monitoring only. Requires dedicated location.
STB / DVR-mining	Automated	None	Dedicated CPE (consumer premises equipment)	Does not track out-of-home and non-TV viewing. MVPD and/or device manufacturer dependent.

Overall, these methods

- Do not address the challenges of multipoint to multipoint viewing environments
- Cannot easily scale across distribution platforms
- Are not based on open, industry standards
- Do not make use of newer, popular consumer media devices (e.g. Smart televisions)

3. Popular New Methods

In order to respond to the needs of a growing, diversifying and more fragmented industry, new strategies must be devised. Ideally, audience measurement must be fully automated, avoid consumer efforts, and save the hassle of rolling out dedicated equipment in mass.

Generally speaking, techniques that rely on a form of digital signature or marker that do not alter the consumer experience and that can be monitored by means of standard, commonly used devices are growing in popularity.

The most common approaches are presented here below.

3.1. Fingerprint

A fingerprint is the digital signature of a piece of content. It is generated at the origination point and need not be transmitted alongside the signal. However, its use assumes the existence of relatively intensive computational capability at the playback point which could be cost or processing or battery-life prohibitive for universal consumer applications. Furthermore, any content alteration whether intentional for distribution purposes or unintentional or malicious - could impact or defeat fingerprint-based content recognition.

3.2. Video Watermark

This method consists of embedding a low visibility digital code in the video signal. These watermarks are very popular for copy protection but have not commonly found their way in audience measurement because, in part, of their decoding requirement that would make them cost prohibitive for that application.

3.3. Audio Watermarking

First experiments seem to demonstrate that Audio watermarking provides most of the benefits sought for audience measurement without the drawbacks or limitations mentioned above for the alternatives. Because of its growing adoption, including by standardization bodies, it is the main subject of this paper.

4. Audio Watermarking Overview

Audio watermarking consists of inserting a signal in the main audio tracks of audiovisual works. The purpose of this embedding is most frequently to identify the piece of content it is implanted in, the medium through which it was transmitted, and, potentially information about its playback mechanism, location, timestamp and viewer.

Applications of this technique range from copy protection to audience measurement. The latter will be detailed further in this document.

From an adoption point of view, two of the leading audience measurement companies (Nielsen and Kantar) are already using audio watermarking for that purpose. But before providing a real-world use-case, general attributes are detailed hereafter.

4.1. Common Workflow

A typical implementation of audio watermarking can be broken down into a few steps illustrated and, then, listed below.

Figure 1 – Workflow Diagram

4.1.1. Watermark Generation

In this step, the actual watermark is created. The watermark must include all information necessary to identify the content and be recognized by detectors at the end of the content supply and playback chain.

4.1.2. Embedder

This stage results in inserting the watermark, typically by means of a specific device, in select audio tracks of the content to be measured.

4.1.3. Content Processing, Storage, Transmission and Playback

These elements are not part of the watermarking process per se. They are simply listed here to help situate the watermarking key steps in the entire content chain.

4.1.4. Watermark detection

The content playback and its contextual parameters mentioned above are registered by detecting the watermark, typically via a functionality embedded in the playback or companion device.

4.1.5. Reporting

The event detection and the contextual parameters are automatically reported to the measurement organization via a connected setup. Here also, the reporting mechanism is typically embedded in an existing apparatus such as the playback or companion device.

4.2. Desired Attributes

In order to address the technical and business objectives of audience measurement, the selected mechanism must present a minimum set of attributes.

4.2.1. Survive the processing chain

The audiovisual content is likely to be edited, reformatted, compressed, transcoded…. before its final restitution. In order to save the content supply chain the burden of changing their workflows, the audio watermark must remain unimpacted by these processes.

4.2.2. Preserve metadata and ancillary signals

The content is increasingly enriched with signals, data and containers (e.g. IMF, SCTE 35, proprietary watermarks...) to enable additional services. The insertion of the audio watermark must be compatible with them.

4.2.3. Easily detectable

The watermark must be detected by low-computing devices that are commonly available to avoid cost or the need for specific devices.

4.2.4. Unambiguous Identification

Finally, and very importantly, the watermark detection must equate to a unique, persistent, error-free, openly-resolvable identification of the content in its exact version (e.g. the accurate episode of a series with the correct edit and encoding level in addition to the time-stamp, network, channel, and/or distribution platform information).

5. Application to Combined Media and Advertisement Measurement

Key industry players came together to define an end-to-end content chain that meets the requirements above. The Coalition for Innovative Media Measurement (CIMM) launched a project called trackable asset cross-platform identification (TAXI Complete) that focused on seamless media identification and measurement.

To that end, the generic workflow depicted in Figure 1 above was translated into a concrete embodiment and enriched with mechanisms for unique identification of media and advertisement contents bound together. This linkage is called open binding of identifiers (OBID).

The system was standardized by the SMPTE Technical Committee for Television and Media (TC-24TB). The standard incorporates the watermarking method as well as the process of embedding standard identifiers for the content and advertisement. Both identifier types are architected to be universal, open, persistent, unambiguous and resolvable.

The entire chain is detailed in industry publications that are summarized hereafter.

TAXI Complete is based on the generation of two watermarks: one to identify the advertisements and entertainment content per the OBID standard, the other to identify the time and channel/distribution platform (or “video service”) per the OBID time label content (OBID-TLC) standard. The watermarks are inaudible and do not interfere with the quality of the program audio signal.

The OBID watermark tandem is usually inserted just before its broadcast but can also be pre-inserted for nonlinear content such as video on demand (VOD).

The OBID content and advertising watermark will typically be accompanied by a timestamp and a video service identifier - collectively called time label content (OBID-TLC).

These OBID-TLC contains up to 4 layers of distributor identifiers to enable tracking the content across a variety of networks. Examples include multi-point, tiered or syndicated distribution networks.

From an implementation point of view, the generation and embedding of watermarks can be integrated into standard pre-existing equipment used by the media production or distribution company. (e.g. the audio processing devices of a broadcast facility).

A functional diagram, based on CIMM publications, is depicted below.

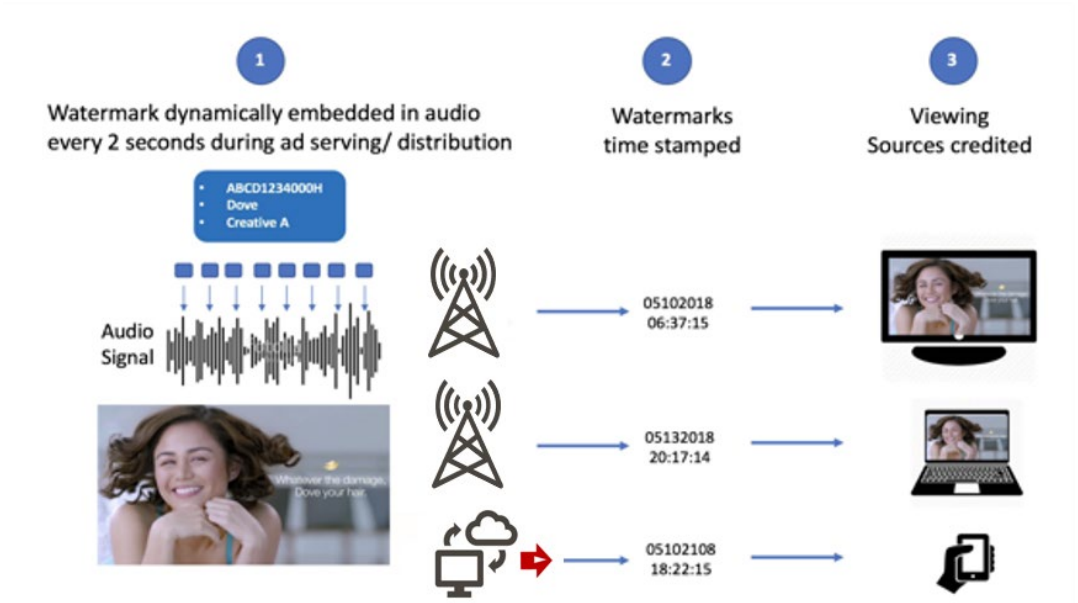


Figure 2 – Typical CIMM Functional Diagram

The watermark tandem is inserted immediately after the creation of the content to allow any content modification downstream without the need to restore the watermarks. Insertions are repeated every 5 seconds for the main content and every 2 seconds for the advertisement.

The detection and reporting are the most commonly used playback devices in particular for broadcasters and MVPDs: typically set top boxes or, as an addition to legacy solutions, Smart televisions.

A complete implementation of the end-to-end system is depicted below:

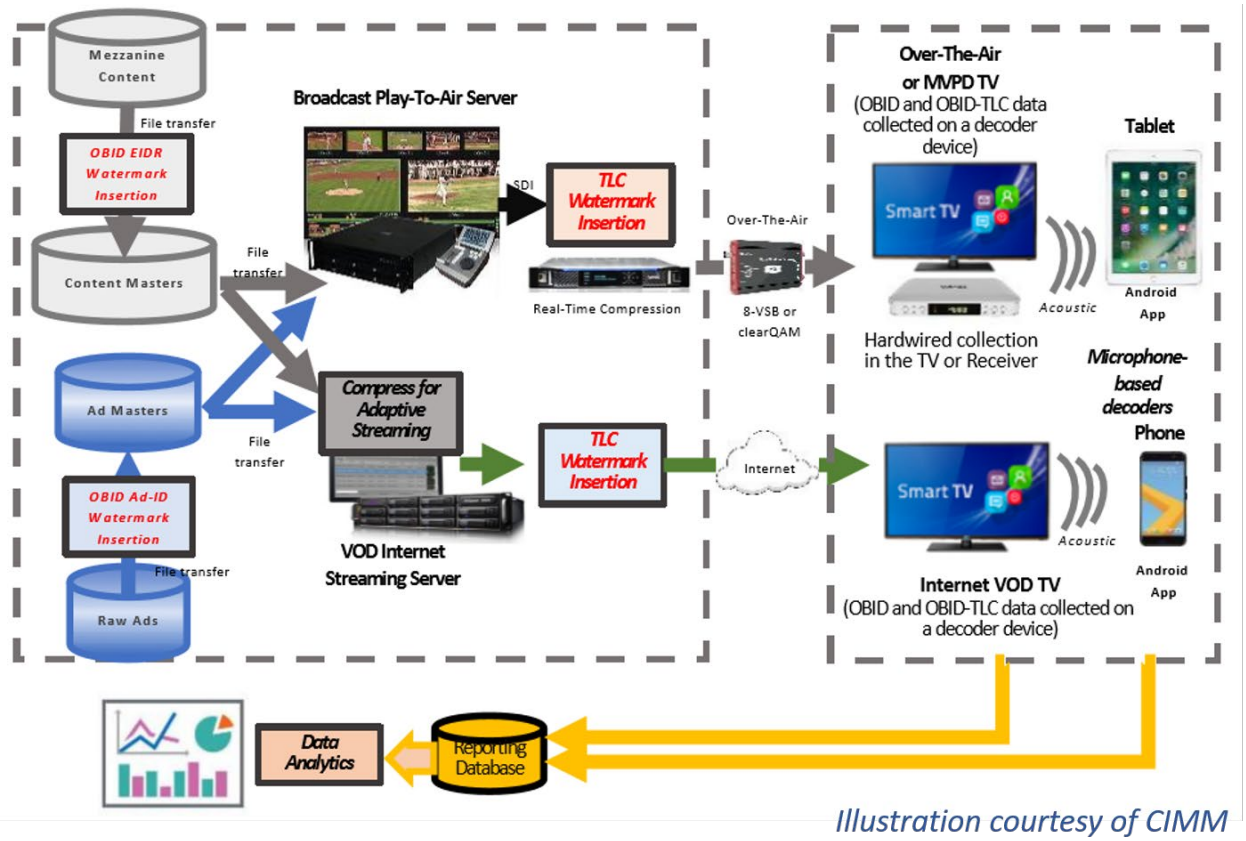


Figure 3 – Typical CIMM Functional Diagram

6. Industry Adoption

The approach has been promoted by multiple industry groups: advertisement identifier (Ad-ID), CIMM and the Entertainment Identifier Registry Association (EIDR), the TAXI Complete OBID method was standardized by SMPTE. It has also received support from an array of broadcasters, MVPDs, studios, advertisement groups, online video companies and more.

It has also entered preliminary testing with the ATSC 3.0 organization and is on the path to become the measurement scheme of reference for this broadcast standard.

7. Laboratory Trials and Results

The field test has been documented by Ad-ID. Edited excerpts of these documents are inserted in this and next sections.

Although the experiments depicted in this section were conducted on an ATSC3.0 network, the technology could equally apply to ATSC1.0 or physical networks used by MVPDs as explained later in this document.

7.1. Test Description

On Friday, 15 February, 2019 a trial of OBID and OBID-TLC standard took place in the Phoenix, Arizona area. The programming and advertising content for the trial (embedded with OBID watermarks carrying EIDR IDs and Ad-IDs) was an encoded version of the CIMM TAXI Complete demo video (previously edited and tested at FOX Television Labs in Los Angeles). The content was broadcast over the air on the trial ATSC 3.0 broadcast channel in Phoenix and was played back on an ATSC 3.0 compatible receiver at the trial site. Representatives from Ad-ID and Pearl TV were present at the Phoenix trial site at that time.

Figure 4 – Video Sequence Tested

The demonstration video had a split screen, where the content is displayed on the left side of the screen, and a video capture of the detector software is shown on the right side, notes on the detection rates were taken as a comparison of the right side of the video.

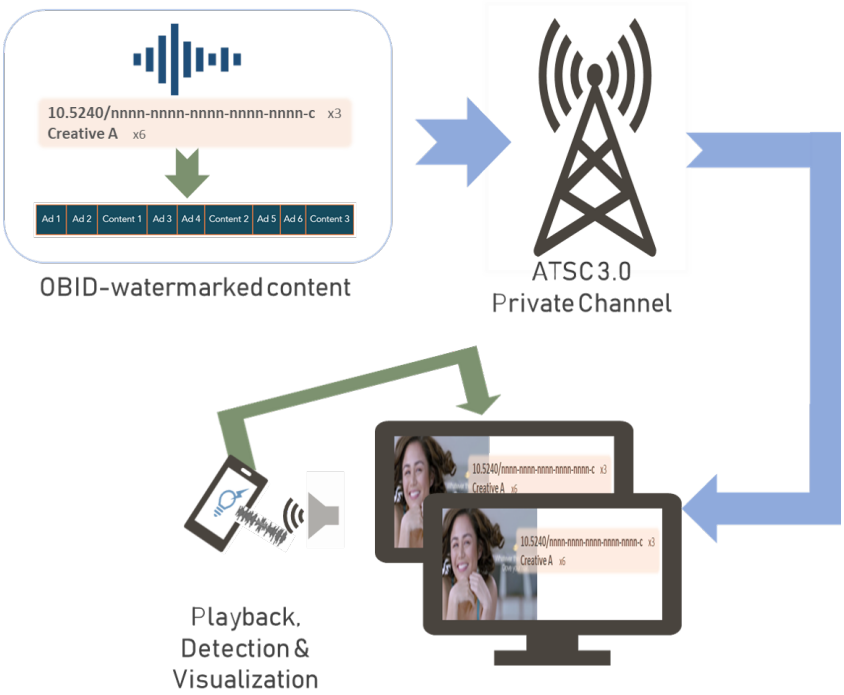


Figure 5 – Test Set Up

7.2. Test Conditions

- The playback devices were common, commercially available, television sets.
- The room where the televisions were located was about 10 ft. by 10 ft., with direct access onto a frequented hallway.
- The hallway door was open during the whole time, to closely replicate potential ambient noise in a real detection environment.
- The detector (a common, commercially available, mobile device running a common commercial operating system) was approximately a foot to a foot and a half away from each television.
- The television volume was at a moderate level, and the audio could be heard at an acceptable level throughout the room.

7.3. Test Results

- At 96 bits, the majority of OBID and OBID-TLC watermarks matched that of the right side of the screen.
- Detections were made of content and ads displayed on two Smart televisions from major consumer brands. Negligible differences were found between the two devices.

8. Limitations Of Watermarking For Over-The-Air-Broadcast

While the acoustic-based trial of CIMM TAXI Complete via ATSC 3.0 transmission in Phoenix was conclusive, one of the challenges of acoustic detection is that it relies on an enabled second / companion device in the home or out-of-home location, other than the initial video playout device.

9. Application To And Opportunities For Cable Networks

The limitations pointed out above for terrestrial network do not apply to cable television. The existence of customer premise equipment such as STB, DVR or gateways in consumer homes enables cable operators to implement detection software without any potential environmental interference (ambient noise) or demand on the user (availability and installation of a companion device or software).

Indeed, STBs and similar devices could excel in being an "all-in-one" content delivery vehicle and data collection device. Content delivery and anonymous content measurement – linked together at the operating system and middleware stack level - with downstream (in-band) and upstream return path data (out-of-band) could be architected to work together.

There are precedents for this kind of “round trip” of viewership data in the MVPD space, from the days of interactive television (in the 2000’s), where STBs were employed using proprietary methods to anonymously measure content viewership (for the purposes of targeted advertising) on a trial basis.

A significant opportunity for investigating the applicability to cable services therefore exists.

10. Conclusions

Audience measurement represents a needed tool for international Broadcasters and the Media & Entertainment industry, as the diversification of sources and distribution channels, and the surge of nonlinear viewing on conventional and new devices, have turned the traditional measurement methods into an incomplete tool.

For that reason, new methods that are non-intrusive, scalable and supply chain-proof are required. Audio watermarking combined with unique identification and open binding point to new generation standards that are advanced and benefit from broad early adoption, driving competition and technological innovations. They can be inserted in the signal in real time or offline, survive editing, encoding and primary, secondary and tertiary transmission.

With OBID, the identification of content anywhere in the distribution chain is possible. OBID-TLC adds the ability to identify how, when, and where content is delivered through the distribution chain to the end consumer. Aside from more accurate tracking and audience measurement of content and ads across platforms, the benefits of such an open standard range from greater efficiencies and cost savings throughout the cross-media ecosystem to improved workflows. After completion of its final test phases (through an iterative process of industry cooperative efforts), it can be rolled-out to successfully measure the next generation of video services.

However, we also see that there are potentially multiple ways to attain the goal. Ubiquitous practice of the necessary standards and systems across the TV industry can only be a net positive shift for the many participants in the television ecosystem.

11. Abbreviations and Definitions

11.1. Abbreviations

Ad-ID	advertisement identifier standard
ATSC	Advanced Television Systems Committee
CIMM	Coalition for Innovative Media Measurement
CPE	consumer premise equipment
DVR	digital video recorders
EIDR	Entertainment Identifier Registry Association
IMF	interoperable master format
M&E	media and entertainment
MVPD	multichannel video programming distributor
OBID	open binding of identifiers
SCTE	Society of Cable Telecommunications Engineers
SMPTE	Society of Motion Picture and Television Engineers
STB	set top box receivers
TC-24TB	Technical Committees for Television and Broadband Media
TLC	time labels to content
VOD	video on demand

11.2. Definitions

Downstream	information flowing from the content provider or network to the user
Upstream	information flowing from the user to the content provider or network

12. Bibliography and References

Digital Video and Social Media Will Drive Entertainment Industry Growth In 2019; https://www.forbes.com/sites/nelsongranados/2018/12/18/digital-video-and-social-media-will-drive-entertainment-industry-growth-in-2019/#752eaa574661

https://www.forbes.com/sites/peterhimler/2014/03/14/niche-media-no-more/#6c2ffffb3cc7

https://www.pwc.com/gx/en/industries/tmt/media/outlook.html

https://www.statista.com/topics/1594/streaming/

https://en.wikipedia.org/wiki/Portable_People_Meter

TAXI Complete Demystified; P. Mears, T. De Kerautem

Taking the Offensive Against Video Piracy

Piracy Models, Anti-Piracy Technologies and Best Practices Continue to Evolve as Awareness Increases

A Technical Paper prepared for SCTE•ISBE by

Steven Hawley
Founder and Managing Director
Piracy Monitor
Advanced Media Strategies LLC
Bonney Lake, WA 98391
steven.hawley@piracymonitor.org
+1 (360) 897 6677

Table of Contents

List of Figures

1. Introduction

Over the past several years, online video piracy has rightfully attained a higher profile in the consciousness of the video industry. The evolution of video analytics platforms used by operators and video providers has been a reflection of this awareness. Pay TV service providers have focused their concerns on reducing the unlicensed use of legitimate services within their service reach, while content providers going direct-to-consumer (OTT) or reaching consumers via online aggregators have focused more on identifying content that is found outside of its legitimate channels of distribution on the Internet.

There also is a growing awareness that the technologies used to identify infringing use, the application of anti-piracy countermeasures, and the decision process that triggers anti-piracy responses must be complemented by well-informed business rules and business policies that guide the selection and use of anti-piracy technology. A complete anti-piracy program also consists of executive level commitment, operational practices and organizational resources that are dedicated to anti-piracy, as well as a network of collaborators in the Internet community and in law enforcement.

Piracy falls into three categories: the theft of content, the theft of services, and theft of advertising. Rather than providing a deeply technical discussion of any particular aspect of video piracy, this article takes three steps back to look at the bigger picture.

2. What is Video Piracy and How Big is the Problem?

Video piracy is the distribution of stolen video content or the redistribution of stolen services, without the rights to do so. Theft can occur as a result of breaches to data centers, video processing and storage, the process of delivery, by breaching the user authentication process, or through capture at the time of playback.

According to a report published by Parks Associates in January 2020, the value of pirate video services accessed by pay TV and non-pay TV consumers may exceed $67 Billion worldwide in 2023.[1] If just 10 percent of pay TV subscribers discontinued pay TV services in favor of video delivered by pirates, the 2023 loss to operators could approach $6 Billion. This is in addition to services revenue lost by pay TV operators due to password sharing. The broader impact of global piracy in the US was estimated to be more than 29 Billion in 2018 alone, by the US Chamber of Commerce. [2]

Even individual piracy cases are quite valuable. For example, in 2019, a piracy operation called Omniverse One World Television,[3] which offered video via a Web portal, through resellers, and via a custom-built illicit streaming device - and even sold advertising - was shut down. In October 2019, Omniverse agreed to pay a $50 Million settlement. In October 2018, SetTV paid damages of more than $90 Million to US satellite TV provider DISH Network and Nagrastar, as the penalty for distributing programming stolen from DISH to more than 180,000 users. SetTV was shut down.[1]

Password (credential) sharing has been seen by the industry as something of a gray area. Some will say that the sharing of credentials outside of the scope of granted *usage* permissions is piracy, even if access is not shared further. Others contend that password sharing isn't piracy unless it is done with the intent to *redistribute* content without the rights to do so. It is not the purpose of this article to settle this question with a legal opinion. In isolated instances, some will even allow infringement as an intentional marketing tactic to boost viewership.

In any case, credential sharing results in significant lost revenue to pay TV operators. A survey of US consumers conducted by Parks Associates earlier in 2019[4] determined that 5% and 6% of those surveyed used someone else's credentials to access pay TV and online video services, respectively. Other estimates are higher.

3. What is Being Pirated and How?

Today, any type of content that can be turned into digital bits is subject to online piracy. Measured in terms of links that are propagated by pirates, more than half of pirated content is television programming, followed by movies (about a fifth of all content), software (about a tenth), games (about a tenth), and published content such as e-books (most of the remainder), according to private research by Piracy Monitor.

If we look at TV programming alone, sports makes up about three quarters of the content delivered via streams that are propagated by pirates. Within the sports genre, football (soccer) is unsurprisingly the most pirated, followed by general sports programming (networks that carry multiple sports), followed by basketball and motorsports. Beyond the sports genre, the most stolen content is TV series and movie programming.

Advertising is also subject to fraudulent use by video pirates, in two ways. The greater threat to advertisers comes from the theft of legitimate advertising by pirates. A CNBC report about a pirate video service called TeaTV [5] described how the service tricked automated advertising services into serving legitimate ads to it programmatically. In such situations, not only is the pirate taking payments under fraudulent pretenses, but also, the fact that the advertiser becomes associated with the pirate may do damage to the advertiser's brand and reputation.

The other form of advertising fraud is by fraudulent video providers, to gain prominent placement within search engine results. Try an online search for "IPTV" and you will see.

4. Vectors for Piracy

The avenue to piracy that has captured the most attention by pay TV operators has been the fraudulent use of end-user credentials, which is essentially a theft of *service*, and not directly the theft of content. To be clear, the act of consumers sharing passwords - or using passwords shared to them by others - is not where most of the credentials used for industrial-scale piracy originate.

The more widespread form of piracy results from the theft of *content*. A variety of methods are available to pirates to capture content, ranging from old-fashioned video camcording in movie theatres and HD television sets, to theft of digital production copies and DVD ripping. Another way is to overcome traditional pay TV conditional access safeguards. A pirate can steal programming at the point of reception by using decoders and stolen keys to decrypt incoming satellite signals.

To identify TV channels with a high likelihood of being stolen, such as a premium pay-per-view event, the video provider can embed invisible forensic watermarks into the video computationally. If the content has been watermarked, pirates can use a process called "collusion" to average a set of multiple instances of the same channel to defeat the watermark before re-encoding it into streaming formats for redistribution.

Large-scale content theft results from the theft of aggregated service credentials. Pirates use consumer databases that were stolen through accidental breaches or the intentional penetration of enterprise data centers and made available for purchase on the Dark Web. These databases may have been stolen from pay TV providers, retailers, financial services institutions, or from other consumer-facing enterprises. Another method is to leverage social media APIs to expose consumer data.

In turn, this consumer data can be used as the basis for phishing attacks upon consumers, in which a pirate masquerading as a legitimate video provider sends a message asking the consumer (for example) to log in and re-set their password. This conveys actual account access to the pirate, which, in turn, opens the potential to steal directly from the video provider's library.

Another way to access video libraries using consumer credentials is to use brute force. Because many consumers use the same user-IDs and passwords for all of their online accounts, credentials from non-media account sources can bear fruit, so pirates will use automation to keep trying account credentials until they find ones that work.

5. Piracy Use-cases: How Pirates Reach Consumers

Once the pirate has acquired the content, video pirates leverage several physical delivery methods:

- "IPTV" streaming[1] which, in terms of the proportion of overall pirate distribution, has increased steadily in recent years. Streaming is more prevalent in North America and Europe
- Direct file download, which is prevalent in Asia, Latin America and Africa and has decreased in recent years
- Peer-to-Peer (torrent) distribution, which is more prevalent in South Asia and Oceania and has also decreased
- Digital lockers, which are file storage services that can host files for download, FTP transfer or torrenting

5.1. Device environments

Consumer environments targeted by pirates include:

- Pay TV set-top boxes where programming is intercepted at output
- Retail streaming devices (such as Roku, Fire TV, et al.), PCs, game consoles, mobile smartphones and tablets, and smart TVs.
- Browsers that reside in PCs and game consoles, accessing pirate streaming Web sites
- Apps that run in legitimate media center environments such as Kodi, which consist of pirated online video and multichannel TV programming. Kodi is available for PCs, game consoles, mobile devices and Raspberry Pi.
- Apps developed by pirates that run in Android and iOS consumer devices, to present pre-linked stolen programming to smartphone and tablet users.

[1] "IPTV" – Traditionally, the term Internet Protocol Television and its acronym IPTV have been used to reference pay TV services delivered through IP multicast over operator-managed networks; originally by the Telcos. It is a supreme irony that this term has been pirated by the pirates, and has become the generally recognized term for pirate video streaming.

- Illicit streaming devices (ISDs) which are custom-designed, produced in quantity, and sold at retail or online. They are preconfigured with an embedded Web browser that is pre-programmed to access pirate video streams. Often, these devices also offer private app stores that allow users to download jailbroken or illicit versions of apps that redistribute legitimate services. Content may be free or at additional cost

Another form of video acquisition by pirates is the capture satellite programming at a receiver, to convert into streams for redistribution.

5.2. Pirate Business Models

Pirates make money by leveraging one or more business models, which include:

- Free-to-consumer model: Some pirates establish Web sites to aggregate pre-programmed links to streaming servers that are hosted by others, into a single user experience that can be accessed from any browser. These sites are often free to the consumer, and funded by fraudulent use of programmatic advertising. Alternatively, they may be funded by revenue shared by providers of ransomware that is surreptitiously distributed via the streaming portal and installed on the consumer's device.
- "Pay TV" subscription model: Some pirates create streaming services with tiered bundles that resemble a pay TV service. Often, these will have "good," "better," and "best" ranges of programming, and will charge a different amount for each programming tier. Revenue comes from monthly payment. The consumer may pay for access using an online payment account or cryptocurrency
- Business-to-Business model: Some pirates assemble turnkey services that are not intended for direct-to-consumer streaming, but rather, host and present stolen content for streaming, direct file download, or P2P (torrent) distribution by resellers. This approach has appeal because multiple resellers amplify the pirate's market presence. The appeal to the reseller is that the reseller does not host any content directly, but rather, acts as a linking site.
- Combination model: Some pirates will do a combination of some or all of the above. One example is Ominverse One World Television, which offered an Android-based ISD with an embedded appstore directly to consumers. It also offered its delivery infrastructure and content to multiple resellers, many of which believed Omniverse to be legitimate. Its programming was a combination of pay TV and online video programming and sold ad insertion space to legitimate multichannel TV advertisers.

Pirates have become increasingly sophisticated and attract consumers with low prices and high production values. One example is in Figure 1 below. In most regions of the world, consumer awareness has been such that most consumers can't discriminate between legitimate and pirate services.

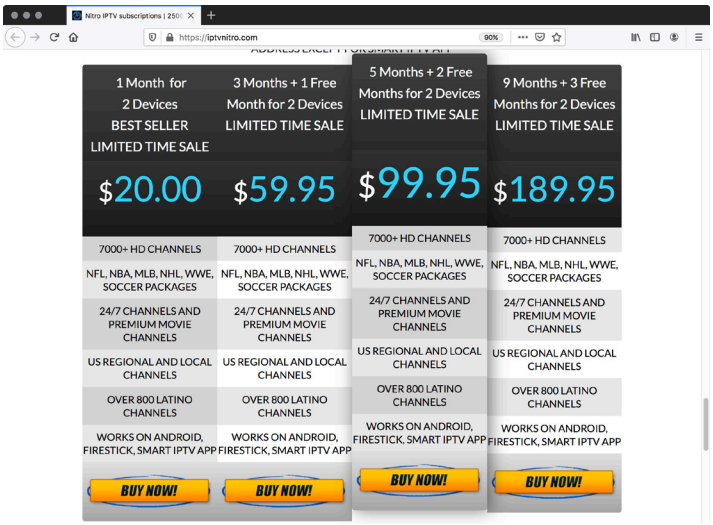


Figure 1: Pirate Video Service with a Tiered Subscription Model (Source: IPTV Nitro)

A comprehensive reference paper about illicit streaming devices, satellite signal re-encoding and other methods used by pirates to steal and deliver video to consumers can be found in the 2018 SCTE-ISBE paper *Analyzing the Modern OTT Piracy Ecosystem*.[6]

5.3. Deepfakes, an Emerging Piracy Use-case

An additional emerging threat is that of deepfakes, which are video clips or programs that are designed to deceive the viewer. Deepfakes can be used to spread disinformation about a given topic or brand by making changes that are relatively easy to make, given today's content production tools.

Stolen content can be embedded within a deepfake video during production. Techniques have also been developed by smartphone providers to create full videos from single still image frames (selfies, for example). The audio track of a video or still image using a trusted spokesperson can be edited or replaced, and the spokesperson's facial expressions can be revised. The fraudulent result may be so good as to be indistinguishable from fact by humans.

Deepfakes can also be produced to attack independent content creators, where images or video content developed by the independent creator can be modified, using an attacker's audio content, for example. The attacker can then approach the independent creator, charge them with theft of its audio track, and extort a ransom for "copyright."

If the source content has been watermarked prior to release, it can be detected if it is used within a deepfake, so the deepfake can be taken out of circulation or its producer prosecuted.

6. Anti-piracy as a Technical Process

Anti-piracy is a combination of detection and action. The process of detecting suspected instances of piracy and then confirming (or dismissing) them as infringing use is automated through the use of monitoring and analytics. Once detected and confirmed, the instance is presented for decision and action.

6.1. Detection and Analysis: The Evolution of Analytics

Video providers have long endeavored to provide high quality service. The pursuit of video quality has changed over time, but the goal has always been in service of the video provider's overall value propositions to consumers. In other words, to make the video experience look and work better.

Video analytics began with the test and monitoring of quality of service (QoS), which is mainly about minimizing delivery errors. This provided a foundation to improve quality of experience (QoE), by measuring the integrity of the content, including video clarity, adherence to color gamut parameters, audio/video sync, captioning, and metadata, and the overall presentation of the experience. Good QoE is dependent upon good QoS.

As video services migrated to Internet Protocol access, it became possible to use the technologies of online ad insertion, ad measurement and streaming quality analytics to better ensure continuity of experience and to measure its effectiveness: advertising analytics.

To detect and address piracy, analytics has taken another step. Video providers can establish usage parameters, watch for usage that falls outside of those parameters, and monitor for content originating from sources that are suspected of piracy to see where that content came from (e.g. was it stolen from your service? From which device and which end user account?). This can be referred to as infringement analytics.

To stay abreast of the constant barrage of monitoring data, automation is necessary to determine whether or not to raise a red flag. Evaluation parameters are typically established by the content rights-holder or content owner. Examples include:

- Number of devices: Allowing account holders to use a defined number of simultaneously active devices. Detecting sudden changes in the range or number of devices associated with an account.
- Allowed devices: to permit delivery to specific types of devices (e.g. HD STBs and streaming devices, but not smartphones).
- Location of use; for example, in-home use only or attempts to access services from unrecognized locations.
- Registered devices: to detect when someone whose device is not registered in a subscriber's household attempts to watch a program – with or without access credentials.
- Anomalous service usage: for example spikes in service access or license requests in a short period of time
- Anomalous content attempts: for example, to make requests through broken or nonexistent links (the equivalent of a Web '404'). Or requests to unrecognized IP address ranges, VPN links, or unrecognized NAT-ted addresses
- …and others

Infringement analytics platforms are also equipped not only to evaluate watermarks, but also to evaluate fingerprints (automated content recognition, video metadata, operator and content provider logo images, and monitor known infringing sites on an ongoing basis.

Figure 2: Analytics for video quality, advertising and rights infringement (Source: Piracy Monitor)

Together, these three approaches - video analytics, advertising analytics and infringement analytics - combine to ensure an overall high-quality experience that conforms to rights parameters.

6.2. A Piracy Decision Loop

Once in place, the technical side of anti-piracy is a process of monitoring, detection, alerting, and then, to apply a desired outcome. There are several methods for doing this.

6.2.1. Credential Monitoring and Analytics

One method is in Figure 3 below, in which an analytics platform is in place to detect the distribution and use of a video service or stream. The platform can detect the user's location, device profile, and other identifying information, to detect infringing use.

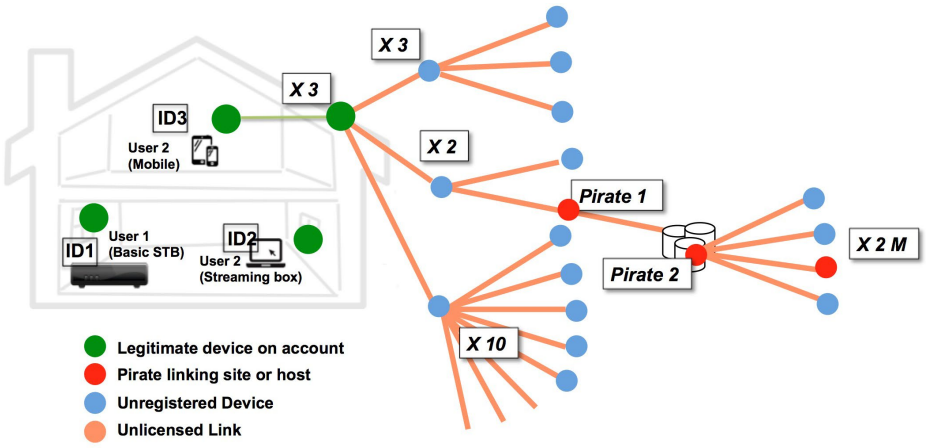


Figure 3: Monitoring to detect out-of-profile service usage (Source: Piracy Monitor)

Users #1 and #2 within the household pictured at left are all legitimate registered users and all of the devices are registered with the service. User #2 is sharing to another receiving device or user within the household. Monitoring and analytics show that this recipient was in turn sharing the service with three other devices or users. One of them, in turn, shared to 3 others. Another shared to 10 others. And while the third user shared to just two others, one of them was a pirate that used that shared access to steal content that was, in turn distributed to a hosting pirate site that served 2 million end users.

Video providers can build 'typical' user profiles for a streaming video account, which then provides a reference point used to detect out-of-profile account usage.

6.2.2. Forensic Watermarking and Monitoring

Rather than monitoring for service abuse, Figure 4 below shows the process of monitoring for stolen content. To begin, the content is prepared for distribution by embedding a forensic watermark. If a video asset is suspected of distribution by a pirate, the video can be evaluated individually for the presence of that particular watermark.

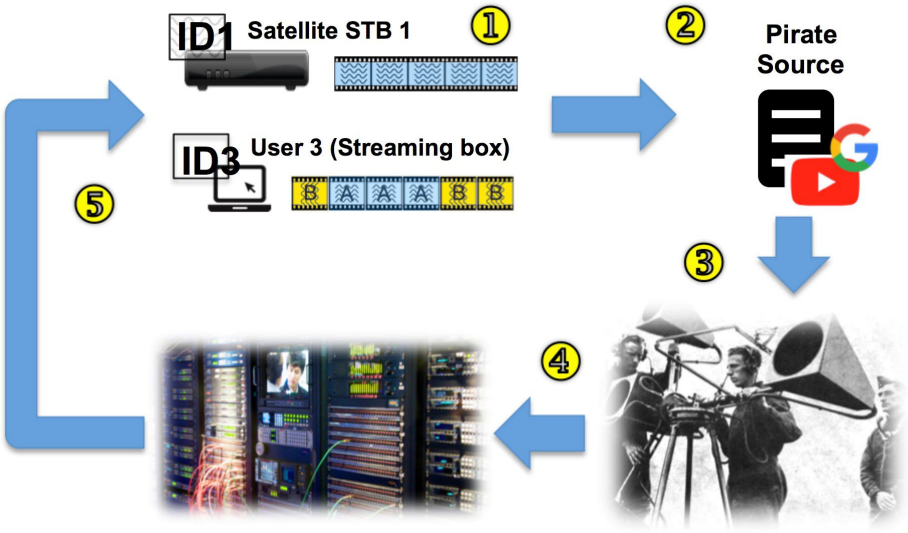


Figure 4: Piracy monitoring, detection, decision and action (Source: Piracy Monitor)

In Step 1 of Figure 4, content is watermarked. This watermark may be embedded at the point of encoding to identify a program that is distributed via broadcast or multicast, as with the video being played by the satellite set-top box. Alternatively, the watermark may be embedded when a streaming session is established, at the service provider's headend or in the CDN. In this example, watermarking is a two step process that first creates duplicate streams with given different watermarks and then segmented. Each streaming session assembles the video segments in a sequence that is unique to that session. Alternatively, watermarks can be applied by a software process running within the streaming client device (not pictured), which eliminates the need for duplicate streams; in turn, reducing the need for storage and processing resources.

In Step 2, a pirate has intercepted the video and is distributing it over the Internet. Step 3 shows the originating video provider monitoring suspected pirate sources for instances of its video content. The monitoring platform has been programmed to alert the video provider (Step 4) that a video is suspected of having been stolen. Upon confirmation, through an automated or a human decision making process, the keys to the device are revoked or the stream is shut down (Step 5).

6.2.3. Automated content recognition

Also known as fingerprinting, automated content recognition (ACR) is in some ways the "inverse" of watermarking: a process used to extract tiny fragments from a video asset without changing the source content itself, and then store these fragments in a database that associates it with an owner or authorized distributor.

Through automation, the video content found on the Internet is compared against the fragments in the database. If it identifies sources that were not licensed to distribute the content, the system signals the legitimate owner or distributor.

6.2.4. Additional Monitoring Techniques

Other anti-piracy approaches include deep packet inspection and evaluation of network flow data. By looking at the request and handshaking process within a video request, a monitoring platform can detect out of range IP addresses, unauthorized virtual addressing, the use of VPNs or proxies, or packet characteristics that may indicate infringement.

Additional reference material is available in past papers published by SCTE-ISBE, including:

- Detecting Video Piracy with Machine Learning (2019)[7]
- Automated Detection for Theft of OTT Services and Content (2017)[8]
- Service Theft in DOCSIS Networks (2019)[9]

7. Anti-piracy: The Long Game

As noted earlier, a complete anti-piracy initiative complements detection and analytics technologies with policies, practices and organizational development. Content protection and anti-piracy technical guidelines are available from multiple sources.

7.1. Elements of an Anti-Piracy Program

Task Force	Establish a dedicated Anti-piracy Team consisting of executive management, with designated technical, financial, marketing and legal experts who are tasked with overseeing, approving and enacting the anti-piracy initiative
Strategy and Goals	Produce an anti-piracy strategy and a set of anti-piracy goals. Develop policies designed to accomplish those goals
Solutions Owner	Empower a program manager to work cross-functionally within the company, to define an anti-piracy initiative and to shepherd it through conceptualization, requirements-development, vendor selection, implementation, operationalization and ongoing improvement.
Risk Assessment	Commission an end-to-end security audit and systems analysis to investigate and confirm the nature and scope of vulnerabilities to piracy. Assess traditional pay TV security and digital rights management as well as IT infrastructure. Consider outside resources with anti-piracy expertise.
Architecture	Establish a reference anti-piracy framework based on risks, goals and policies, informed by the risk assessment, and by technical and financial feasibility analysis. Recommend a first choice and a fallback approach from among multiple possible solutions.
Resources	Determine the program elements, resources and enabling technologies that best fulfill goals and policies.

Operations	Develop a dedicated operations resource responsible for piracy, with a reporting process, a severity-ranking system, and escalation and resolution procedures. Consider building a simulation environment to replicate attacks and to evaluate alternative solutions.
Cybersecurity	Because many piracy risks exist outside of video processing and delivery, identify points in data centers and in the cloud where content, personally identifiable information and internal resources could be exposed to exploitation or compromise
Intelligence	Evaluate emerging piracy, anti-piracy and cybersecurity use-cases on an ongoing basis, to continually improve those practices. Stay abreast of regulation that may affect you.
Community	Join organizations from the media, entertainment and technology industries that focus on piracy in your region. Sign up for their infringement and piracy alerts. Establish relationships with search engines and other online providers that may have contact with your content or services.
Law Enforcement	Gain an understanding of the governmental and regulatory agencies with jurisdiction in your own territory, as well as for those in regional and global markets (such as the European Union); and for international law enforcement agencies such as Interpol. Identify your local liaison officers.

It is also important to recognize that the piracy problem is not static. Pirates and hackers are very creative and live in a culture where smart people challenge other smart people to create more effective traps. Some of them are individuals that operate in the dark, while others are nation-state actors. Some of them eventually decide that their energies are better used to join the fight against piracy.

7.2. Technical Guidelines

In addition to the MovieLabs recommendations noted earlier, several other media industry organizations publish guidelines for content protection and antipiracy. They include:

- The *Enhanced Content Protection Specification*, from MovieLabs [10]
- *Content Protection Best Practices*, from the Motion Picture Association (MPA/MPAA) [11]
- The *Ultra HD Forum Guidelines*, [12] from the Ultra HD Forum
- *Forensic Watermarking Implementation Considerations for Streaming Media*, from the Streaming Video Alliance [13]

8. Conclusions

In today's streaming video world, piracy is a fact of life. The US Chamber of Commerce estimated the impact of global online piracy to the US economy in 2018 to be more than $29 Billion in lost revenue. According to a 2019 study released by Deloitte,[14] 2018 marked the first year that streaming video captured more consumers than traditional TV.

While credential theft and account abuse have been the focus of pay TV operators and online content aggregators, content owners and producers have been more concerned with theft of the content itself, no

matter whether it is the result of theft of OTT services or of a breach to delivery infrastructure or consumer devices.

High-value content is more likely to be stolen, where value is measured by exclusivity, choice, immediacy, age and quality.

Exclusivity: Programming that is exclusive to a single programmer is more likely to be stolen. Consider *The Mandalorian*, the *Star Wars* series that was introduced with the launch of Disney+ by The Walt Disney Company in the US, Canada and the Netherlands in November 2019. Streams were detected by Google Trends in many other countries almost immediately after Disney+ went live.

Choice: With the emergence of so many SVOD streaming services, consumers have to choose. It is becoming more and more likely that any given consumer will want *something* that is not available via the services that they subscribe to, so the consumer might attempt to access it from a pirate source rather than pay for yet another SVOD service.

Immediacy: Pay-per-view live sports programming is most valuable when a match or a game is in its early stages. This makes it incumbent on sports programmers and video providers that carry that programming to be in a position to detect illegal streams, isolate their sources, and take action in minutes.

Age: The age of the content is critically important. Just as media companies stage the distribution of their content in different release windows for different distribution channels, with the most valuable windows coming first; new releases are more likely to be pirated

Quality: Because ultra high-definition programming has become more mainstream, UHD resolution has become less of a differentiator and therefore is less likely to justify a higher fee just because it is UHD. However, UHD quality also enables a pirate to generate high quality streams that nullifies any differentiation based on quality by legitimate online video providers. This is why Movielabs, MPAA and others have issued formal guidelines for forensic watermarking, to be able to detect stolen UHD content and make it easier to take it out of circulation.

Digital piracy as we know it today arguably had its origins nearly 20 years ago with the emergence of music hosting sites like Napster and KaZaA, which hosted stolen audio content for download using peer-to-peer protocols. In retrospect, Apple's iTunes service was probably the legitimate alternative that had the most impact in reducing music piracy while enabling content owners and rights-holders recapture some of the revenue that had been lost to pirates.

For video, a similarly disruptive 'silver bullet' solution has yet to emerge.

9. Bibliography and References

[1] Hawley, Riney, Kent. *Video Piracy: Ecosystem, Risks and Impact*. Research report. Parks Associates. January 2020. See: https://www.parksassociates.com

[2] Blackburn, Eisenach, Harrison. *Impacts of Digital Video Piracy on the US Economy*. Research report. NERA Economic Consulting and the US Chamber of Commerce Global Innovation Policy Center. June 2019. See: https://www.theglobalipcenter.com/wp-content/uploads/2019/06/Digital-Video-Piracy.pdf

[3] *Complaint for Copyright Infringement. Demand for Jury Trial*. Legal complaint against Omniverse One World Television, a pirate operation. United States District Court. Central District of California. Western Division. February 14, 2019. See: https://www.documentcloud.org/documents/5740024-Omniverse.html

[4] Naden, Jiang, Kamble. *360 Deep Dive: Account Sharing and Digital Piracy*. Research report. Parks Associates. July 2019. See: https://www.parksassociates.com/blog/article/pr-07162019

[5] Megan Graham. *Netflix and HBO shows are getting pirated on this app that's been bankrolled by advertisers such as Pandora, BET+ and TikTok*. Article. CNBC. October 20, 2019. See: https://www.cnbc.com/2019/10/20/netflix-and-hbo-shows-are-getting-pirated-on-teatv-and-other-sites.html

[6] Jones (Comcast), Foo (Charter). *Analyzing the Modern OTT Piracy Ecosystem (2018)*. White Paper. SCTE-ISBE. September 2018. See: https://www.nctatechnicalpapers.com/Paper/2018/2018-analyzing-the-modern-ott-piracy-video-ecosystem

[7] Tooley, Belford. *Detecting Video Piracy With Machine Learning (2019)*. White Paper. NCTA. September 2019. See: https://www.nctatechnicalpapers.com/Paper/2019/2019-detecting-video-piracy-with-machine-learning

[8] Catranis, Yuan, Belt (Irdeto). *Automated Detection for Theft of OTT Services and Content (2017)*. White Paper. NCTA. 2017. See: https://www.nctatechnicalpapers.com/Paper/2017/2017-automated-detection-for-theft-of-ott-services-and-content

[9] Westervelt, Florendo, Belt (Irdeto). *Service Theft in DOCSIS Networks (2017)*. White Paper. NCTA. 2017. See: https://www.nctatechnicalpapers.com/Paper/2017/2017-service-theft-in-docsis-networks

[10] *Movielabs Specification for Enhanced Content Protection*. Technical Guidelines. Movielabs. Revised 2018. See: https://movielabs.com/solutions-specifications/enhanced-content-protection-ecp/

[11] *Content Protection Best Practices*. Technical guidelines document. Motion Picture Association. October 2019. See: https://www.motionpictures.org/what-we-do/advancing-creativity/additional-resources/#content-protection-best-practices

[12] *Ultra HD Forum Guidelines*. Technical guidelines document. Ultra HD Forum. September 2019. See: https://ultrahdforum.org/guidelines/

[13] Stevenson (Ericsson), Wilkenson (Comcast). *Forensic Watermarking Implementation Considerations for Streaming Media.* Techical guidelines document. Streaming Video Alliance. July 2018. See: https://www.streamingvideoalliance.org/books/forensic-watermarking-implementation-considerations-for-streaming-media/

[14] *Digital Media Trends Survey, 13th Edition*. Report. Deloitte. March 2019. See: https://www2.deloitte.com/us/en/insights/industry/technology/digital-media-trends-consumption-habits-survey/summary.html

Blockchain based Verification Method for Alternate Content Switching and Dynamic Advertising

A Technical Paper prepared for SCTE•ISBE by

Srilal M Weerasinghe PhD
Principal Engineer, Charter Communications
srilal.weera@charter.com
720-699-5079

Table of Contents

List of Figures

List of Tables

1. Introduction

Blockchain applications in digital TV advertising are generally in the realm of supply chain management. This is understandable given its distributed ledger framework for recording transactions. In this paper, novel applications of the blockchain technology in digital TV advertising and alternate content switching are presented. The impetus for this is the new video auditing requirements of the Society of Cable Telecommunications Engineers SCTE 224 (2018) standard [1], or more commonly known as Event Scheduling and Notification Interface (ESNI/'es-nee').

Sports blackouts are a familiar occurrence on traditional TV networks. During the blackout event, the content distributor is contractually obligated to block the regular content stream. The proof of 'switching to alternate content', is sent back to the Programmer post-event. In traditional cable TV, the blackouts were generally geo-location (zip/postal code) based. Thus, the verification data were a lot simpler for IRD-based (integrated receiver/decoder-based) sports blackouts.

With the advent of web-based digital TV however, a new form of blackouts is now in vogue. This stems from the rights restrictions for content distribution on the web. The new SCTE 224 standard offers rich capabilities to support a wide variety of alternate content and advertising scenarios. As the programmers introduce new features, the content distributors are expected to support them. However, it is also known in the industry that the validation of content switching in IP streaming is a formidable challenge. Sending out a large number of customer device data raises privacy concerns as well. The paper describes a blockchain-based solution to address this intractable issue.

Another objective of the paper is to present how blockchain technology can be utilized in ad verification. The motivation for this is the Interactive Advertising Bureau's 2018 Video Ad Serving Template (VAST) 4.1 standard [2], which is now superseded with 4.2 version. The crypto-hash based solution presented here would constitute as proof that an ad was displayed as contracted.

2. Exploring Blockchain Technology for Digital Advertising

One of the striking aspects of the blockchain model is the duality of transparency and anonymity. This interplay between seemingly opposite concepts can be explored in dynamic advertising. For example, in today's complex digital advertising ecosystem, there are multiple intermediaries involved: Ad agencies, Media Buying Desks, Aggregators, Demand Side Platforms (DSPs), Supply Side Platforms (SSPs), Ad exchanges, Ad networks, Yield optimizers etc. The 'Transparency' feature of Blockchain could reveal the transactional data at a more granular level. This is a boon to advertisers, who are not always clear about the murky details of intermediary transactions. The transaction markups at each stage are somewhat opaque and has been a concern. Understanding the markups at each stage of ad buying/selling process could help combat ad-fraud and lower the cost of advertising [3]. Similarly, the return path from publisher to advertiser is also mired in hard to verify metrics on ad viewability. Since a blockchain is a public ledger for transactions, it can essentially record how many times an ad was viewed. Advertisers and publishers expect that blockchain will bring transparency into the ad buying process [4].

While the ‘Transparency’ feature is heralded as a panacea for all kinds of problems, the ‘Anonymity’ aspect of blockchain has received a bad rap. This is mainly due to the ‘untraceable’ nature of the cryptocurrency bitcoin. Its use in nefarious activities in the past is widely known and has raised a few eyebrows.

In contrast, digital advertising proposes a more innocuous use of the anonymity feature. Blockchain based products can potentially safeguard consumer personally identifiable information (PII) in targeted advertising. Blockchains are decentralized, peer-to-peer networks with no central authority. All transaction records are encrypted with asynchronous keys (public/private) to ensure privacy and security. Anonymity is thus a critical feature, as it would enable trustworthy interactions between parties who may or may not know each other. The application to digital advertising is in addressing a common dilemma that network operators face today. That is, how to share consumer statistical data without revealing the customer PII. The driver is the delivery of more granular targeted ads for higher revenue margins. The risk is any leak of customer PII would have disastrous consequences.

Security of a blockchain is based on the ‘hash tree’ concept known as Merkle trees. Each transaction in the blockchain has a hash associated with it. The hashing process is continued in a hierarchical manner, (as an inverted tree). The result is a root node hash that could verify the integrity of all the transactions. Any tampering with the data at any level of the tree would not only modify the hash value of the tainted transaction, but that discrepancy would be propagated all the way to the top of the tree (‘Merkle root’).

On the regulatory aspect of consumer privacy, there are sweeping changes happening in Europe. The European Data Protection Regulation (GDPR) recently introduced strict laws to protect consumer digital privacy. Companies must now adhere to stringent compliance requirements and safeguard consumer personal data or face hefty fines. Similar laws would soon be enacted in the US, such as the California Consumer Privacy Act (CCPA) bill. It would thus be prudent to examine the privacy aspects of Blockchain model for targeted advertising.

3. Standards Bodies Activities

3.1. SCTE 224

SCTE 224 ESNI is a xml-based protocol that enables the transmission of event and policy information in IP video delivery networks for rights management. Primary use case is for content providers/programmers to communicate upcoming schedules or signal-based events and corresponding policy to content distributors. Application scenarios are blackouts, alternate content switching (ACS) and ad break opportunities.

Defined in the standard are the primary entities Media, Policy, Viewing Policy and Audience. The standard delineates for each **Media** (Channel) what video source is to be used for a given **Audience.** In this context, ‘Audience’ may include the Geo Location and Device Type at any given time. **A**udience characteristics and viewing policies associated with each audience are communicated via XML messages. Policies are key to SCTE 224 implementation. **Policies** are applied or removed by **MediaPoints**. Individual events (MediaPoints) are defined using start and end times (MatchTime) or in-band signaling (MatchSignal) data.

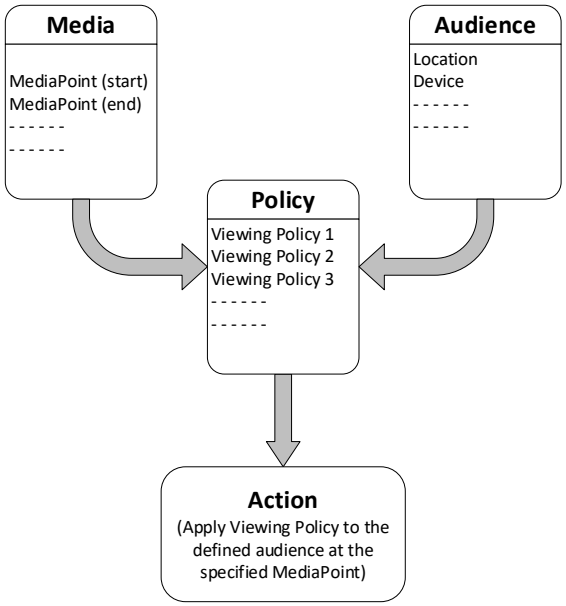


Figure 1 - SCTE 224 Implementation Model

The SCTE 224 standard enables rich metadata—and pertinent to present discussion—an Audit element (Section 8.9 of SCTE 224-2018). Auditing data include status information. Examples are: audience member inclusions, status of ViewingPolicy applications to Audience/ Media, state changes as well as system errors.

In the IP-based video delivery, content distributors need to ensure 'rights management' by tracking Media, Audience, Policy and ViewingPolicy elements. Central to this enforcement is the actual proof of content switching. Needless to say, this is a huge amount of audit data that content distributors are required to supply to programmers. We envision that blockchain, with its indelible recording feature for maintaining a ledger is well-suited for this purpose.

3.2. IAB VAST 4.1

Developed by the Interactive Advertising Bureau, VAST 4.x is the industry standard for communication requirements between ad servers and video players. The new version includes improved ad verification and measurement methodologies that was previously done using the VPAID standard. This is a welcome change as the intended purpose of VPAID was ad interaction and not verification. The new elements, '<AdVerifications>' and '<ViewableImpression>' of VAST 4.x would enable publishers to verify and track ad viewability on their inventory.

To complement the above described standards-based efforts, a blockchain-based ad and ACS verification methodology is outlined in the paper.

4. Use Cases

4.1. Alternate content switching

Sports blackouts are a familiar occurrence on traditional TV networks. During the blackout event, the content distributor is contractually obligated to block the regular content stream. The proof of 'alternate content switching' (i.e. ACS verification data), is sent back to the Programmer post-event. In the web-based digital TV/media delivery, a new form of blackouts is now in vogue. This stems from the rights restrictions for content distribution on the web.

Examples are:

1) Golden Globe awards viewership may be limited to in-home devices only.

2) An old black and white movie may not have rights yet cleared to be viewed on tablet devices.

3) A 4K movie may only be shown on certain iOS/Android devices due to version incompatibilities.

In each case, a subset of viewers would be precluded from watching the scheduled program and would be directed to a static image ('Slate') or to another TV channel (alternate content).

The new SCTE 224 standard offers rich capabilities to support a wide variety of alternate content and advertising scenarios. As the programmers introduce new features, the content distributors are expected to support them. Additionally, the content distributors are required to report back to the ACS compliance results to content providers. However, it is also known in the industry that the validation of alternate content switching in IP streaming is a formidable challenge [5]. Sending out a large number of customer device data to 3rd parties also raises privacy concerns. De-identifying the data is an option, but susceptible to re-identification [6].

4.2. Ad Verification

'Ad fraud' is a major concern for the industry as evident by recent news[6][6]. The ad tech industry is fighting back on several fronts: on the enterprise sector, new initiatives are on the rise, such as applying blockchain technology to improve transparency [7]. In parallel, the industry consortiums are developing new standards to provide tools for the operators to combat ad fraud.

To complement these efforts, we present a modified infrastructure for blockchain-based ad verification. Note that the supply-chain-based advertising eco-systems (with DSP, SSP, Ad-Exchange components) is not the primary mode for some major enterprises. A common practice is to use dedicated ADS services more than open-exchange bidding. In both cases though, the challenge is to provide valid proof that an ad was displayed as contracted.

5. Solution Overview

Blockchain usages offers an indelible record of transaction data. The integrity of the data is maintained by cryptographic hashing of each block as well as tying it to the prior block. Hashing (such as SHA-256), is a one-way function that is extremely hard to break. It is not possible to reverse engineer a secure hash and obtain individual data components. Even a minute modification to the data will affect the hash value of the entire block, as well as the chain.

We propose the usage of 'events' in place of 'transactions' in the proposed blockchain application. The hash-tree would be built based on the ACS event records received from end-user devices. Each such record will contain data about the ACS event as shown below. These data are in a structured format (such as JSON), so that the calculated hash is unique per record.

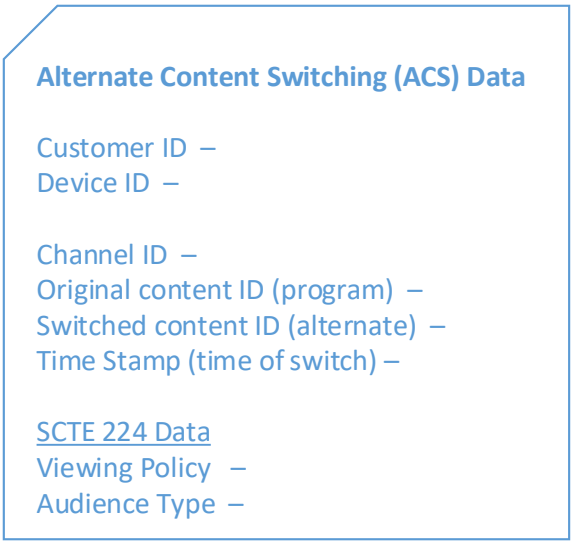


Figure 2-Structure of an ACS Data Record

Next, each record is cryptographically hashed in a hierarchical manner ('Merkle tree') until the root-hash is reached. The final root-hash would constitute proof of the recorded events in the block, since it is mathematically impossible to reverse the process. Having the final hash (as sufficient verification of ACS contractual obligation), would also negate the need for sending a vast amount of 'sensitive' customer data back to each Programmer. Thus, the proposed solution enhances customer privacy.

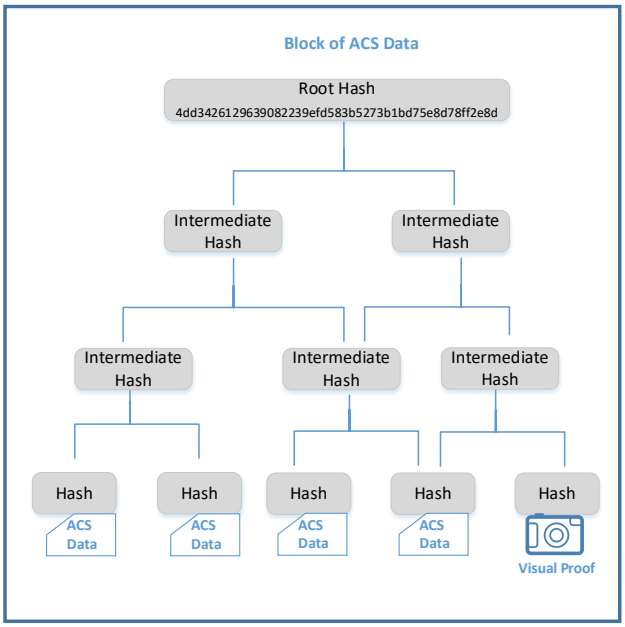


Figure 3-Blockchain ACS data block Structure

5.1. Visual Proof Hashing

Current practice is to supply voluminous data about content switching back to Programmers. Releasing sensitive customer data is a privacy risk. The proposed solution addresses this issue via the crypto-hashing technique. Adding a video segment or audio clip containing a timestamp of the event occurrence would fortify the argument that the hash is sufficient proof. Note that it is infeasible to modify/add the video clip later and obtain the exact same hash (Figure 3). This provides additional visual proof of the content switch. The video capture entity can be located at a data center or cable network headend/hub-office. Note that in other contexts such as online media, 'screen capture' may substitute for video proof.

A requirement for the video hashing as envisioned here is the ability to endure changes to format and speed. A technique currently in vogue is 'perceptual hashing'.

5.2. Perceptual Hashing

In cryptographic hashing, small changes in the input would drastically change the output. While such a characteristic (known as 'avalanche effect') is essential to secure data integrity, it is not suitable in the current application. The reason being that videos are subjected to different speeds, formats and configurations. Each such transformation would produce a different hash, rendering the process ineffective. Similarly, image processing such as cropping and resizing would generate entirely different hashes.

Unlike cryptographic hashing, perceptual hash is computed on a reduced quality image (resized, gray-scaled and averaged over pixel intensities). The resulting hash will not change even if the image is subjected to any transformation (as long the content integrity is preserved). Thus, perceptual hashing is

resilient to variations in aspect ratio, color, skewness, scaling etc. Minor changes in input would not cause major changes of the hash. Put another way, similar images would have similar hash values.

In the proposed solution, the perceptual hash of the image/video is included in the cryptographic hash computation of the block. Note that combining the video hash with dissimilar data in the Merkle tree calculation to generate proof of ACS is a novel concept.

5.3. Creating the Blockchain with ACS/Ad Data

Each block is formed with 'event' data (as in Figure 3). Blocks are connected via hashes to form the chain as in Figure 4. There are different ways to configure the chain. For example, if there are 5000 participating devices in the ACS, each block may contain 1000 units and linked to form a blockchain with 5 blocks, as shown. Alternately, each block may contain ACS data specific for a geographic region (spatial distribution), or a ViewingPolicy restriction applied through a period of time (temporal distribution).

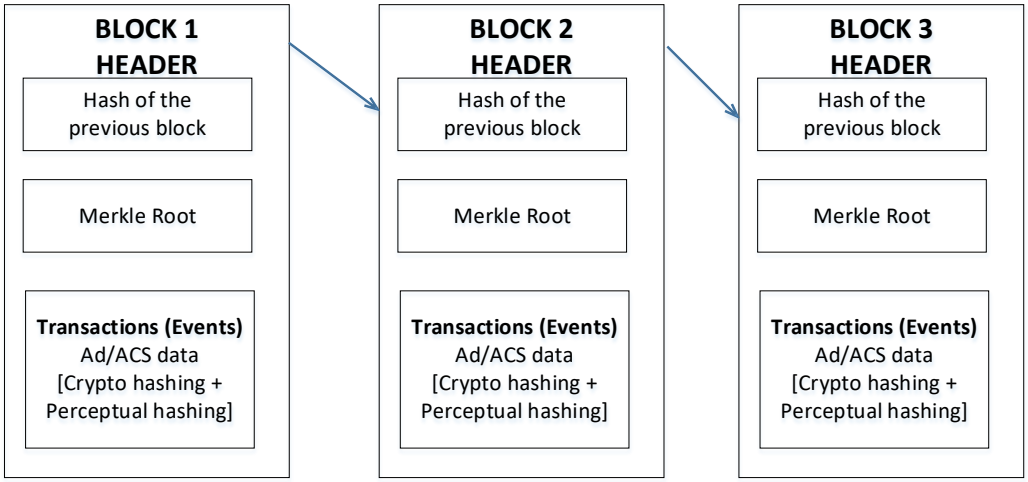


Figure 4-Structure of the Blockchain for AD/ACS Validation

5.4. Proposed Solution Compared to Other Blockchain Implementations

Blockchain architecture varies by implementation and it is hard to define a standard model. For example, some cryptocurrencies do not have the 'distributed ledger' feature. Oracle and Smart Contract concepts were introduced by Ethereum and were not part of original Bitcoin protocol. Also, unlike Bitcoin, most private blockchains don't have group consensus mechanism (mining). And some even advocate centralized control.

In spite of the ambiguity, all major blockchain implementations share one commonality: that is automated crypto-hashing-based block formation, which are then linked to form the blockchain. The proposed solution shares that property as well. But it does not use coins and there is no buying/selling among participants. The proposed solution only utilizes cryptographic hashing to record 'ACS events'. The hierarchical hashing is used to stamp/identify each data block and connect as a linked-list to form a blockchain. A representative video or audio clip is included with the ACS data for visual proof.

Table 1-Blockchain Feature Comparison with Proposed Solution

Blockchain Feature	**Regular Blockchains** (Ethereum, Bitcoin)	**Proposed Solution**	**Comments**
Blockchain type	Mostly Public	Private	
Cryptographic hashing	Y	Y	
Transactions	Buying/Selling	Event (Content switch)	
Block creation	Y (Transaction bundling)	Y (bundling of event data into blocks)	* (see below)
Blocks formed into chains	Y	Y	
Block Mining	Y	N	No mining. No 'Nonce'.
Consensus Mechanism	Byzantine Fault Tolerance	N/A	** (see below)
Smart contracts	Y	N	***(optional)

*Root-hash of Merkle tree is computed. The root-hash is then combined with previous block hash of the chain to compute block hash, As this is a private blockchain with full control by the network administrator, the concept of 'data mining' (or the usage of 'Nonce' to adjust the difficulty), is not relevant. Any manifestation of Byzantine faults will be resolved by the network administrator, who has full ownership of the chain.

** Since there is no interaction among the devices (except with the server), consensus mechanisms such as Proof-of-Work (POW) are not applicable. Technically, the proposed solution falls under POS (proof-of-stake), with one party governing the decision-making process.

*** Smart Contracts are code snippets with conditions and actions listed. They run on top of the blockchain network layer. In general implementations, they trigger payments once the conditions of a transaction are met. In the disclosure, video events are used in place of transactions.

6. Solution Architecture

In this paper we articulate a blockchain-based method for ACS and ad display event verification. In the sections below the architecture and process steps for each scenario are presented.

6.1. Blockchain based ACS Verification

For each ACS policy implementation, the Programmer will receive sample data as shown below.

Table 2-ACS Policy Execution Example

ACS Policy (Hypothetical scenario)	ACS Execution Results	Proof of Validation
Emmy Awards (7 – 9 PM) viewership to be limited for in-home devices	**Mobile** devices that tuned into Channel 123 during 7 – 9 PM = 99% of devices were content restricted (ACS success rate = 99%)	Hash#
4K content to be blocked on iOS-7 (or Android-4) or earlier versions	Apple devices running older iOS that were tuned into 4K content, during the last 30-day period	Hash#

The Content Provider shall consider the 'hash value' as sufficient proof of ACS. This is in lieu of receiving a large number of individual customer data. For example, a reported data could be, 'we have achieved 99% compliance, and here is the root-hash # as proof'. Only in an audit session (conducted jointly by programmer/distributor in a controlled environment), the actual customer data would need to be revealed/reviewed.

In one implementation, each subtending device records the content-switch event (transaction). The ACS event data is passed to the blockchain server which collates them into blocks (Figure 3), and chains (Figure 4).

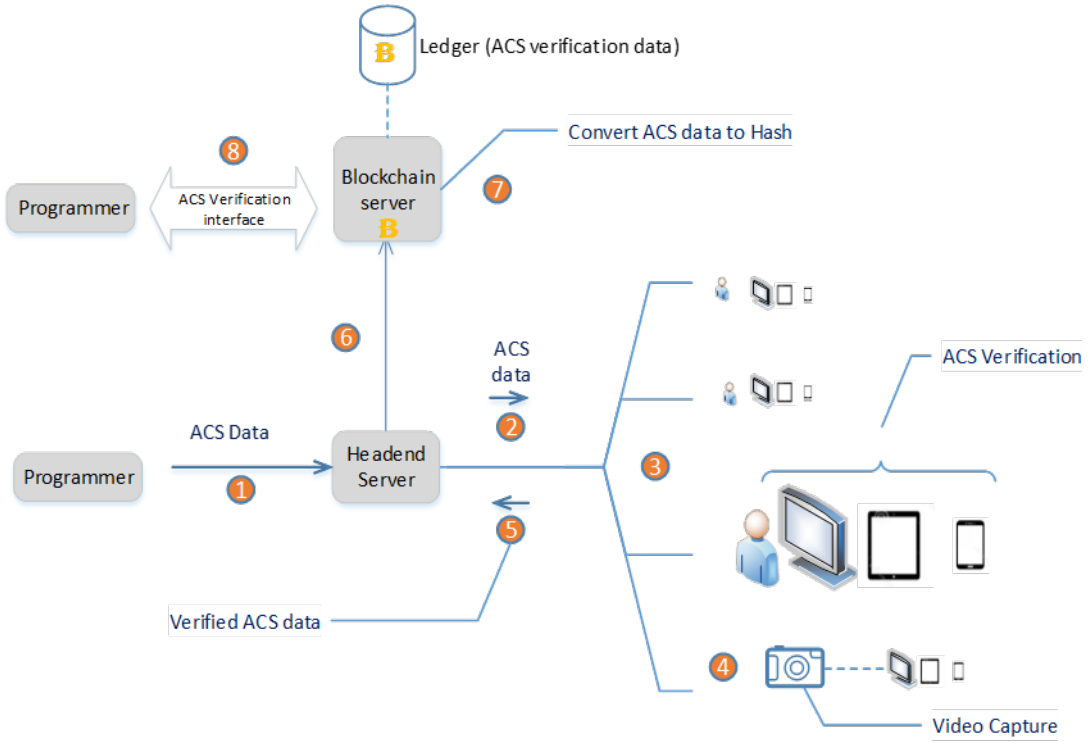


Figure 5-ACS Verification

6.1.1. Process Steps

1. ACS data are received from multiple programmers. A simplified view is presented here for the purposes of the discussion.
2. ACS is applied to multiple devices. The devices could be end-user/customer devices or test devices located in a video operations data center, or on a virtual machine environment.
3. Upon verifying the occurrence of an alternate content switch, the ACS data are recorded per Figure 2. The visual verification could be done manually by an operator, or via an automated program such as those based on machine learning.
4. The 'video capture' is an additional metadata created during the ACS event. If the ACS event did not occur, the failure is captured and recorded as well.
5. The composite data is collected by the headend server (via a data push/pull).
6. Headend server transfers the ACS data stream to the blockchain sever.
7. At the blockchain server, the data is hashed and formed into blocks and chains. The blockchain ledger is saved in a repository.
8. ACS verification data (hash) is sent to the Programmer. Also, an interface (based on PKE) is created for accessing the data by programmers (or proxies).

6.2. Blockchain based Ad Verification

The architecture and process steps are parallel to the ACS case presented above. Instead of programmers sending ACS data via 224, multiple advertisers send ad creatives to the ad server. In the process steps, substitute 'advertisers' in place of 'programmers' and 'ads' instead of ACS events.

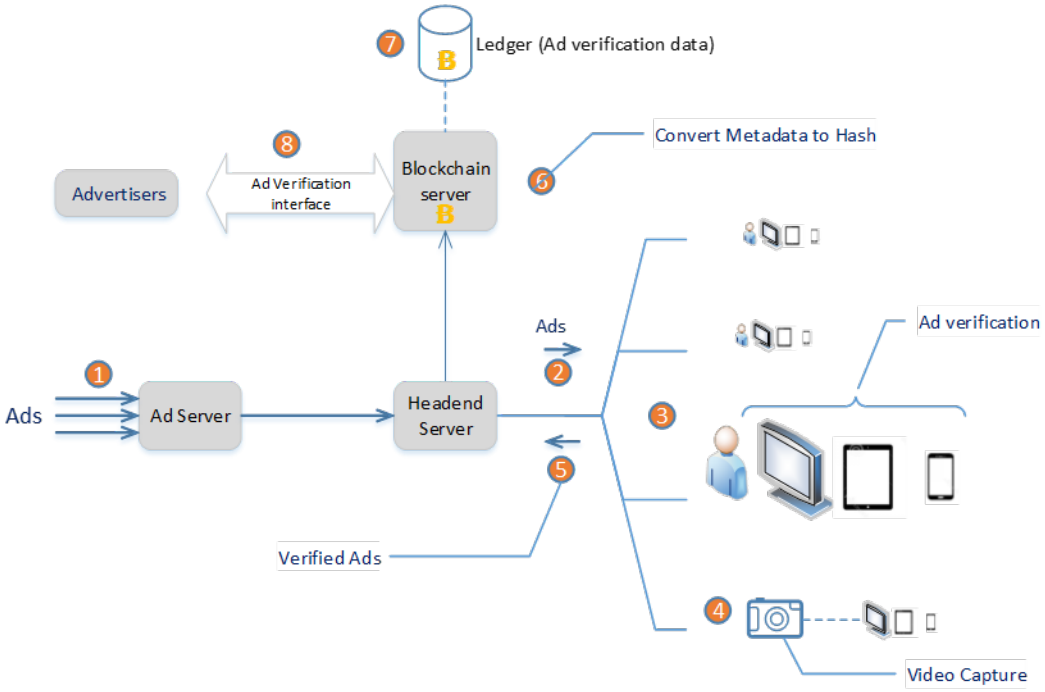


Figure 6-Ad Verification

6.3. Visual Proof Using Machine Learning

The use of deep learning-based techniques such as Convolutional Neural Networks to detect scene changes in videos is an active research field [8]. In the blackouts case, auto-identifying the slate can be done using machine learning. Similarly, in ad detection or content switch scenarios, heuristics-based machine learning analysis can be employed to auto-detect scene changes. The content capture of the video clip/image would be used as input to construct the visual proof hash.

6.4. Distributed Architecture Based Solution

In this section we present a distributed network-based solution. First, a few definitions on **Smart Contracts** and **Oracles**. Blockchain records data sequentially, thus data points from outside the chain require an intermediary to access the chain. This function is performed by an Oracle. Once the specified conditions are met, the Oracle would trigger a smart contract, (containing the instructions in computer code).

Unlike in regular blockchains there are no financial transactions scoped in the proposed solution. Instead, ACS and ad events would trigger the Oracle to initiate the code snippet for writing events to the block. Then the blockchain protocol takes over, cascading the tree formation and the creation of blockchain.

An architecture common to Ad service and ACS is depicted in Figure 7. The 'mesh network' consists of servers running the Blockchain protocol. Multiple devices subtend from Oracles (with multiple Oracles subtending from Blockchain servers). Whenever 'events' occur, devices report those events (via a standard push/pull mechanism) to the blockchain. The network elects a 'compute-node' per event stream as the designated creator of blocks. The selection could be based on the rules previously established such as the current load. In this scenario, other nodes would forward event data they receive to the compute-node. We do not see the need to replicate the chain to all nodes (no mining). Blockchain is also tasked with validating the transactions (events). This includes integrity of data, that no event is reported twice, and additionally a heartbeat signal to ensure the devices are not down.

The calculated hash data are stored in an external server (interface-sever) database for transmittal /retrieval by external entities, such as Programmers or Advertisers.

We envision the blockchain network to be internal to the enterprise (behind the firewall). Alternately, it could be shared by multiple enterprises via VPN tunnels (a federated/consortium blockchain).

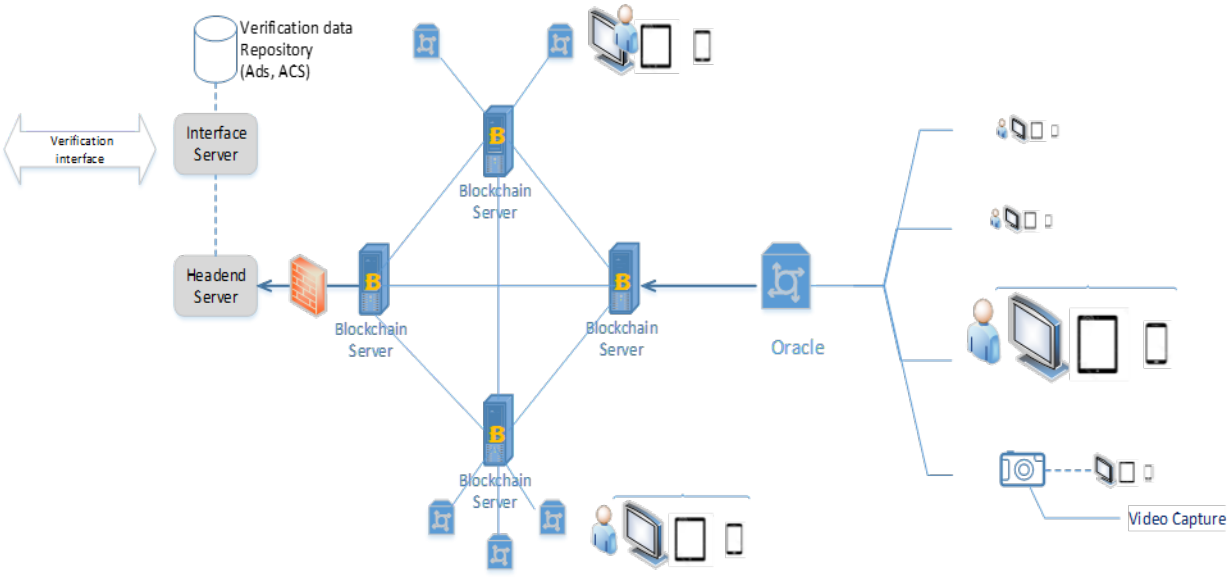


Figure 7- Consortium Blockchain for Ad/ACS verification

7. Conclusions

The recent advances in standards bodies have facilitated new service opportunities in digital adverting and ACS. In this paper a crypto-hash based verification method is presented to supplement the framework set by the standards. Block hash calculations (perceptual and cryptographic) are made with composite data derived from textual and audio-visual sources.

The beauty of the blockchain-based solution is that it negates the need to send voluminous data from content distributor to programmer as proof of 'content switching'. Just the blockchain hash would suffice, as the data is unalterable. Only in an audit session (conducted later jointly by programmer/distributor in a controlled environment), would the actual customer data need to be revealed/reviewed.

The proposed solution preserves consumer privacy while supporting next generation of ACS and digital advertising services.

8. Abbreviations

ABR	adaptive bit rate (a streaming technology via the web)
ACS	alternate content switching - supplanting one video stream (default) by another (alternate)
ADS	Ad Decision Server
CDN	Content Delivery Network
Content Provider /Programmer	content owners, such as major networks and studios
Content Distributor/Affiliate	Network operators that distribute content. (MVPDs and V-MVPDs)
DSP	Demand Side Platform (represent buyer/Advertiser)
IAB	Industrial Advertising Bureau
JSON	Java Script Object Notation
MVPD	Multi-Channel Video Programming Distributor
V-MVPD	Virtual Multi-Channel Video Programming Distributor
PII	personally identifiable information
PKE	Public Key Encryption (generates cryptographic asymmetric key pairs for security)
SSP	Supply Side Platform (represents seller/Publisher)
SCTE 224	Society of Cable Telecom Engineers ESNI standard
VAST	Video Ad Serving Template (IAB Standard)

9. Bibliography and References

[1] SCTE, "Event Scheduling and Notification Interface-ANSI/SCTE 224 2018r1" (2018).

[2] IAB, "Video Ad Serving Template (VAST) version 4.1" (Nov 2018)

[3] Jeff Stone, "The FBI is diving deeper into the Methbot ad fraud case", CyberScoop – https://www.cyberscoop.com/methbot-ad-fraud-fbi-white-ops/

[4] IAB, "Blockchain for Video Advertising" (Feb 2018)

[5] 4-part overview of SCTE 224 – https://www.youtube.com/watch?v=OnViVlsuuCo

[6] Guardian, "New York taxi details can be extracted from anonymized data" – https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn

[7] Forbes "Industries And Activities That Could Benefit From Blockchain's Transparency" – https://www.forbes.com/sites/forbestechcouncil/2019/06/25/13-industries-and-activities-that-could-benefit-from-blockchains-transparency/#7991888fb0fd

[8] SPIE Digital Library, "Deep learning-based scene-awareness" – https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging/volume-28/issue-1/013038/Deep-learning-based-scene-awareness-approach-for-intelligent-change-detection/10.1117/1.JEI.28.1.013038.short

Encoding Intelligence for Optimal Viewer Experience in Live Video Distribution

Letter to the Editor prepared for SCTE•ISBE by

Zhou Wang, Professor, University of Waterloo
Chief Science Officer, SSIMWAVE Inc.
200 University Ave W, Dept. of ECE, University of Waterloo
Waterloo, Ontario, N2L 3G1, Canada
zhou.wang@uwaterloo.ca
519-888-4567 ex. 35301

Abdul Rehman, CEO, SSIMWAVE Inc.
375 Hagey Boulevard, Suite 310
Waterloo, Ontario, N2L 6R5, Canada
abdul.rehman@ssimwave.com
519-489-2688

Kai Zeng, Lead Researcher, SSIMWAVE Inc.
375 Hagey Boulevard, Suite 310
Waterloo, Ontario, N2L 6R5, Canada
kai.zeng@ssimwave.com
519-489-2688

1. Introduction

Real-world live video distribution systems are often faced with the great challenge of processing videos of extremely diverse content type and complexity. The challenge becomes even greater given the critical real-time requirement and the large volume of 24/7 video streams that need to be processed. Using a fixed encoding setup to drive the live video encoders for bandwidth reduction, as is the case in most real-world live distribution systems, causes serious problems, resulting in encoded/transcoded videos that often suffer from severe and unpredictable quality variations across time, video assets, and content types.

In the case of live video distribution, decisions need to be made instantaneously to make the best options for encoder configurations easily adopted in the video encoding/transcoding pipeline.

To empower the encoder with intelligence requires two key components:

1. A quality-of-experience (QoE) metric that not only accurately predicts end viewers experience when consuming videos streamed to their viewing devices, but is also real-time and light-weight, producing consistent QoE predictions across content type, content complexity, codec type, bit rate, video resolution, frame rate and dynamic range; and

2. An intelligent optimization engine that drives the encoders to produce the best and controllable QoE scores in diverse environment and meanwhile maximizing bandwidth reduction.

Working solutions that best address these critical issues are highly desirable for live video distributions.

2. User Experience Metrics for Encoding Performance

An objective user QoE metric aims to automatically predict end viewer's visual experience when watching the encoded video fully decoded and rendered on their viewing devices. Objective QoE assessment is a difficult task because it requires deep understanding about how the sophisticated encoding process creates compression artifacts for diverse types of video content and how such artifacts impact the quality assessment behavior of the human visual system (HVS). Traditionally a direct numerical measure, namely the peak signal-to-noise-ratio (PSNR), has been commonly used for encoder evaluation and comparison, but PSNR has been shown to have low correlation with perceived video quality [1]. There has been a great deal of effort in the past two decades developing advanced objective metrics that better predict subjective video quality. Representative metrics include the structural similarity index (SSIM) [1], [2], the multi-scale SSIM (MS-SSIM) [3], the information content-weighted SSIM (IW-SSIM) [4], the video quality model (VQM) [5] and the video multi-method assessment fusion (VMAF) [6]. These metrics demonstrate significantly improved video quality predictions under certain controlled test conditions. Nevertheless, they are still highly limited in terms of their functionality, interpretability, application scope, and computational cost. Such limitations often make it extremely difficult, if not completely impossible, to use these objective metrics in various real-world video encoding/transcoding scenarios, especially in time-critical applications such as live video distributions. In recent years, novel objective QoE metrics designed to overcome these problems are emerging. These metrics target two types of crucial properties, which will be elaborated here.

The first type of properties focus on the accuracy, speed, cost and interpretability of the QoE metric. There is no doubt that the QoE metric should produce video quality scores that accurately predict viewer experiences. The standard way to test the accuracy of an objective metric is to compute the linear correlation coefficient, rank-order correlation coefficient, and mean prediction error, between the objective scores and

mean subjective opinions using large-scale subject-rated video databases. The metric also needs to have low computational and implementation cost, readily deployed in large-scale video distribution systems. This will also allow for high-speed computation for continuous 24/7 real-time assessment of high-resolution, high frame rate and high dynamic range videos with moderate hardware configurations. The metric must also be easily interpretable, producing quality scores that linearly relate to what an average viewer would say about the quality of a video. For example, if the quality score range may be between 0 and 100, divided into five evenly spaced segments corresponding to five perceptual QoE levels of bad (0-20), poor (21-40), fair (41-60), good (61-80), and excellent (81-100) quality, respectively. Such a metric creates an easy-to-grasp common language, allowing smooth communication in large organizations, where engineers and operators can identify and fix quality problems on the fly, researchers and developers can optimize individual components and the overall video delivery systems, and executives can make critical business decisions.

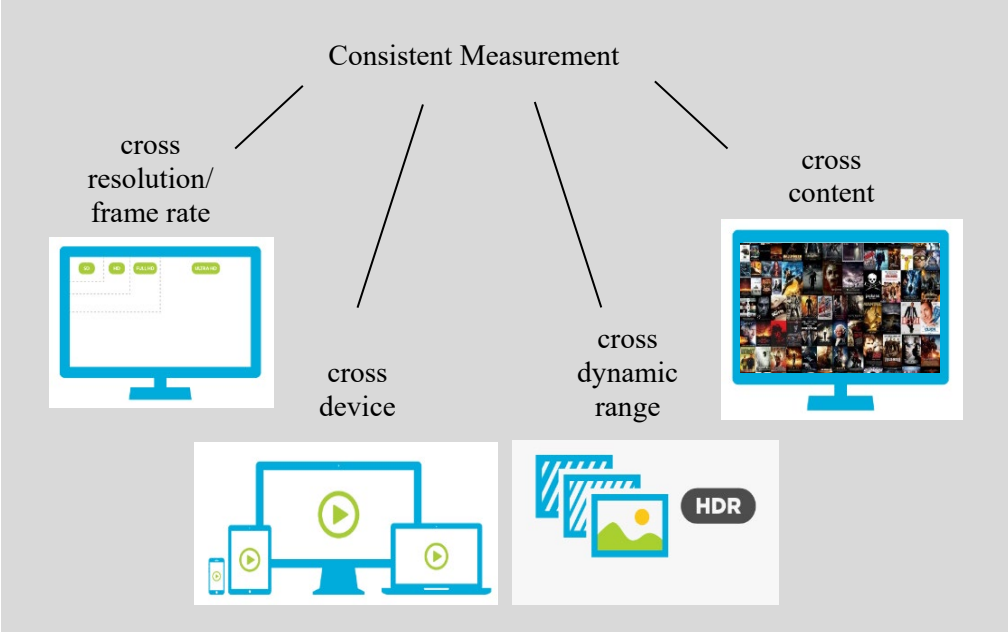


Figure 1 – Critical requirements lacking in traditional QoE metrics

The second type of critical properties relate to the usability and consistency of the QoE metric in real-world application scenarios. It is important to note that well-known video quality metrics (PSNR, SSIM, MS-SSIM, IW-SSIM, VQM, VMAF) require pixel-to-pixel correspondence between the reference and test videos. As a consequence, when videos at the input and output of the video encoder/transcoder are of different spatial resolutions, frame rates, and dynamic ranges, these metrics often do not apply. This greatly impedes the practical usage of these metrics because in modern video distribution, it is very common that the source input videos are transcoded into multiple versions of not only different bit rates, but also different spatial resolutions, frame rates and dynamic ranges. In addition, the playbacks of the same video stream on different viewing devices could create significantly different viewer experiences, but these metrics often generate one quality score only (or a few scores corresponding to a few different devices), and thus fail to capture the device variations of visual QoE assessment. Another common but important issue with these quality metrics is that they often create inconsistent scores across content of different types and complexity

levels. As a result, scores generated by these metrics cannot be compared across content, meaning that two videos of similar perceptual QoE may be given drastically different scores, largely constraining the practical use such QoE metrics in large-scale distribution systems that make instantaneous resource allocation decisions across hundreds or thousands of video services and live video channels. Therefore, as shown in Figure 1, in real-world video distribution systems, it is essential to use a QoE metric that simultaneously produces consistent quality measurements across spatial resolutions, frame rates, dynamic ranges, viewing devices, and video content.

Recently, great effort has been made to develop novel QoE metrics for the above-mentioned properties. So far, the full SSIMPLUS Viewer Score metric is offering all these critical properties [7],[8], and the open source VMAF project has also been making progress towards the direction [9].

3. Encoding Intelligence Driven by User Experience Metrics

A good QoE metric that satisfies the critical properties is a fundamental ingredient to enable encoding intelligence. On top of that, an encoding decision-making engine driven by content and encoding performance analysis may be used to control the live encoding/transcoding process. This may be done in different ways, and two types of encoding intelligence frameworks are described below.

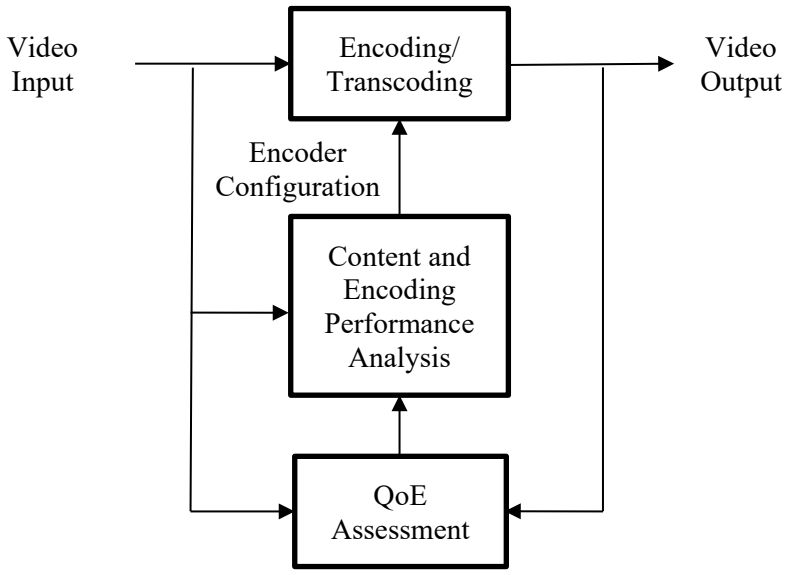


Figure 2 – Type I Encoding Intelligence

The first type of encoding intelligence works for the application scenarios where the encoder or transcoder configurations can be controlled on-the-fly. These configurations may include the spatial and temporal resolutions, the bit rate, the quantization parameter (QP), the group-of-picture (GoP) structure, the encoding pre-set, and other parameters that may influence the encoding process. When the source input video is received, it first goes through content analysis that may include spatial, temporal and color complexity measurement, content type analysis, dynamic range and color statistics, and other statistics of the content. Meanwhile, the QoE metric, which compares the current input and output video streams before and after the encoder/transcoder, is computed and then fed into the analysis module. Based on both content and

encoding performance analysis, decisions on encoder/transcoder configurations are made and used to control the encoder/transcoder instantaneously. The intelligence decisions should be geared towards the best balancing point between sustained quality delivery and cost-effective bandwidth usage. This process is illustrated in Figure 2.

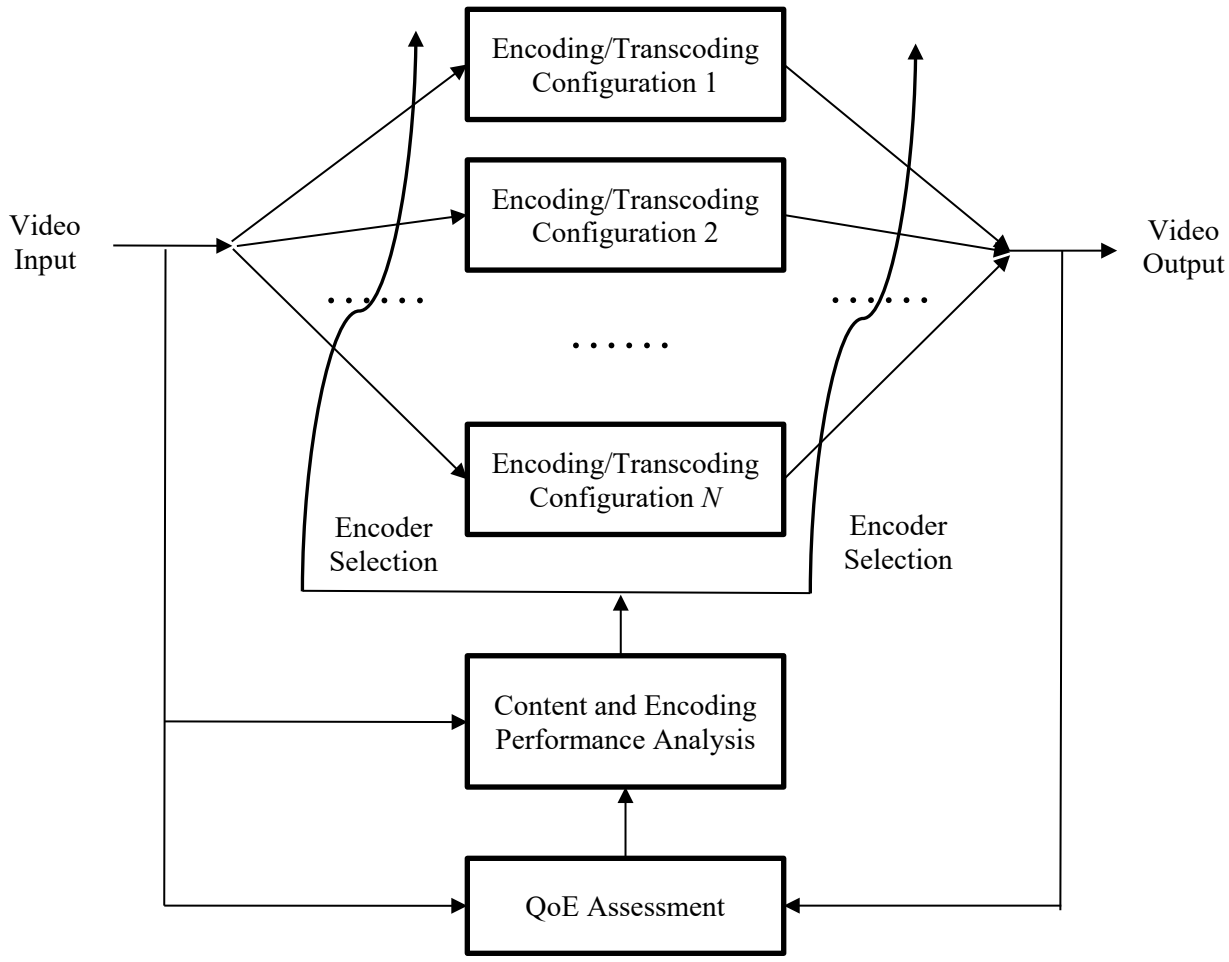


Figure 3 – Type II Encoding Intelligence

The second type of encoding intelligence adapts to the scenarios where on-the-fly encoder parameter adjustment is difficult, but multiple encoder/transcoder configurations are setup previously. As a result, the intelligence is on the selection of encoders from multiple options, as shown in Figure 3. The pre-determined encoder/transcoder configurations may be designed to target at videos of different content types and spatial/temporal/color complexity levels. They could also represent different types of encoding technologies or encoder solutions. Similar to the Type I intelligence case, source content analysis is performed and the QoE metric between the current input and output video streams before and after the encoder/transcoder is computed instantaneously. Both types of information is employed by the content and encoding performance analysis module to create an intelligence decision that choses one out of the multiple encoder/transcoder configuration options for the next step or encoding event.

In both encoding intelligence frameworks, each encoder/transcoder block may be designed to generate one output video stream or a ladder of outputs (which includes multiple encoded videos of different resolutions, frame rates, and bit rates), depending on the deployment points in the video delivery chain and also on the specific use cases. In addition, the analysis and decision-making processes may be based on either short-term instantaneous inputs, or on long-term statistics.

4. Conclusions

Compared with video-on-demand (VoD) and many other use cases, encoding intelligence for live video distribution is more challenging because all the critical decisions need to be made instantaneously, any suboptimal decisions need to be identified and corrected on-the-fly, and the solutions need to work robustly and continuously 24/7 in large-scale systems. The tolerance of errors is often low, and any wrong decision may lead to severe and unpredictable quality issues, immediately affecting a large number of end viewers' visual experiences [10]. The two most crucial components for encoding intelligence is the QoE metric and the encoding intelligence engine. We discussed the challenges and state-of-the-art solutions for both components. We have also discussed two types of general frameworks on how QoE-driven encoding intelligence may be deployed in real-world application scenarios.

5. Bibliography and References

[1] *Image quality assessment: from error visibility to structural similarity*, Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, IEEE Transactions on Image Processing, Apr. 2004.

[2] *Video quality assessment based on structural distortion measurement*, Z. Wang, L. Lu, and A. C. Bovik, *Signal Processing: Image Communication*, Feb. 2004.

[3] *Multi-scale structural similarity for image quality assessment*, Z. Wang, E. P. Simoncelli and A. C. Bovik, IEEE Asilomar Conference on Signals, Systems and Computers, Nov. 2003.

[4] *Information content weighting for perceptual image quality assessment*, Z. Wang and Q. Li, IEEE Transactions on Image Processing, May 2011.

[5] *A new standardized method for objectively measuring video quality*, M. H. Pinson, IEEE Transactions on Broadcasting, Sept. 2004.

[6] *Toward a practical perceptual video quality metric*, Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy and M. Manohara, Netflix TechBlog, Jun 2016.

[7] *Display device-adapted video quality-of-experience assessment*, A. Rehman, K. Zeng and Z. Wang, IS&T/SPIE Electronic Imaging: Human Vision & Electronic Imaging, Feb. 2015.

[8] *SSIMPLUS: The most accurate video quality measure*, https://www.ssimwave.com/from-the-experts/ssimplus-the-most-accurate-video-quality-measure/

[9] *VMAF: the journey continues*. Z. Li. C. Bampis, J. Novak, A. Aaron, K. Swanson, A. Moorthy and J. De Coc, Netflix TechBlog, 2019.

[10] *Begin with the end in mind: a unified end-to-end quality-of-experience monitoring, optimization and management framework*, Z. Wang and A. Rehman, SMPTE Motion Imaging Journal, 2019.

SCTE • ISBE

Society of Cable Telecommunications Engineers
International Society of Broadband Experts

SCTE :: Society of Cable Telecommunications Engineers
ISBE :: International Society of Broadband Experts

140 Philips Road | Exton, PA 19341-1318 | T: 800.542.5040 | F: 610.884.7237 | scte.org • isbe.org