Automated Detection for Theft of OTT Services and Content

Identifying Your Content Out in the Wild

A Technical Paper prepared for SCTE/ISBE by

Lucas Catranis
Product Marketing Director
Irdeto
Hoofddorp, NL
lucas.catranis@irdeto.com

Brian Yuan
Product Owner
Irdeto
Ottawa, Ontario CA
brian.yuan@irdeto.com

Dave Belt
Technology Evangelist
Irdeto
Conifer, Colorado US
303 653-7647
dave.belt@irdeto.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

Introduction

Theft of video content has been an issue since the dawn of pay television (TV). Since that dawn there has been a continual security cat and mouse game between operators and pirates, with the technology actively evolving with it. Most of this technology has, however, focused primarily on protecting the video pipeline with the assumption that control ends at the playback device.

Within this paper we look at content beyond the device, after it has been played back, pirated and distributed over the Internet. We'll first look at the common methods of obtaining pirated content; then, we'll look at some of the state of the art techniques be used to combat these.

Indeed, the next generation of content protection must move beyond the device to understand and mitigate the path of active online piracy.

The Problem

As mentioned, theft of premium video content has and is a continual problem. Loss of content revenue undermines the business models of all participants of the video production and delivery pipeline. This theft is now reaching a new unprecedented level due to the ubiquity of the Internet coupled with the ready availability of tools, devices, and piracy services available to the layman. The piracy services themselves are now becoming active and imminent competitive threats to the operator's business model, many looking like legitimate over-the-top (OTT) services to the novice.

Content itself has a natural decay with regards to its value over time. In general when we speak of "premium content" we are frequently referring to the age of the content itself or the quality of the encoding and delivery. As such, live sporting events, early release movies and ultra-high-definition (UHD) content, are classified as premium content. This content has sufficient demand for a pirate to successfully monetize and also serves as the greatest loss to the content owners and operators.

Many of the techniques discussed herein are aimed primarily at premium content due to its value to the providers.

Content Theft Methods

1. Multicast

In multicast systems, a hacker can go through the exercise of compromising the set-top box (STB) via Joint Test Action Group (JTAG) interface hacking, side channel attacks and the like. If one is merely interested in obtaining high quality video though, the easiest route is via the high-bandwidth digital content protection (HDCP) port.

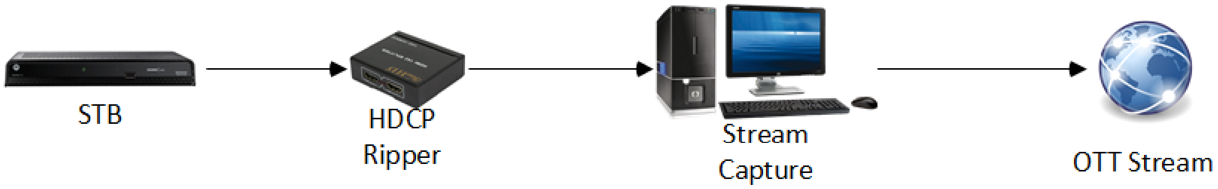

Figure 1 – Workflow using HDCP Ripper

As shown in Figure 1, an HDCP ripper in plugged into the high-definition multimedia interface (HDMI) port of an STB. The ripper simulates a valid HDCP client and is able to decrypt the content delivered over the port. The decrypted stream is then captured via a personal computer (PC) and re-broadcast over a streaming service.

The ease with which HDCP rippers are obtained and the simplicity of such a system make online piracy detection the only mitigation to such a scenario.

2. Over the Top

Theft of OTT services offers a different set of opportunities. If one has a 10’ viewing device in the form of an internet protocol (IP) STB, one can again use the HDCP ripper described above. Other attacks involve the sharing of credentials which are necessary for authentication on unmanaged devices. Hacking of an operator’s application itself can enable free service to a large population via distribution of the application.

While we do our best to secure the delivery of the content up to the viewing screen, content can and is getting out. To plug this hole in the revenue model we need to secure content from the outside in.

Content Identification Methods

Due to the prevalence and dynamic nature of content piracy, correctly identifying content in a timely and cost effective fashion is key to significantly reducing the amount of content theft which leads to subscriber churn. Various methods are deployed to identify pirated content with levels of information varying based on the technique. To get a highly accurate picture of the theft of a particular piece of content, multiple complementary methods should be deployed.

1. Metadata Analysis

The first step in identifying content is to comprehensively sort through the myriad of content freely available on the internet to the content which is specifically infringing on the content provided by the service provider. Step one typically uses keywords and metadata analysis in a variety of search modes; from search engines to social media, the internet should be scoured for potential streaming/content downloading sources. Common techniques include multi lingual keywords (both positive and negative), metadata analysis for validated users, and metadata analysis based on posted data such as file name, type, and run time. Depending on the content life cycle, these keywords are aggressively applied by automated systems to rapidly identify potential new targets of interest.

2. Video Fingerprinting

Well architected metadata analysis still typically generates hundreds to thousands of potential real time infringements for each broadcast event. Sorting through each of these with human analysts is neither timely nor cost effective, so the solution is to use automated video identification techniques such as video fingerprinting.

Video fingerprinting typically takes a real time or video on-demand (VoD) file for an upcoming event and analysis of the video for unique characteristics to create a unique digital signature of the upcoming event. This digital signature is, in turn, compared automatically against the gamut of infringements found via metadata analysis and sifted into three categories: full match, partial match, and no match.

The key to effective fingerprinting is the accuracy of the fingerprinting technique. For example, there are many common video obfuscation techniques employed to circumvent the fingerprinting techniques of common sites such as YouTube. And while sites like Facebook employ highly effective (proprietary) implementations of audio fingerprinting, the current state of their video fingerprinting technology is still quite easy for content pirates to circumvent. For proper content identification, a layered identification approach is recommended to ensure that content is correctly characterized and identified in a timely fashion.

3. Deep Learning Image Recognition

There is a great deal of information which can be learned from piracy data; source identification, subscriber demand, cryptographic integrity, and source leakage are just a few of the pieces of information which can be obtained by carefully analyzing piracy data. As an example, the logo can tell you which broadcast infrastructure sourced the leak. This information, in turn, can tell you the efficacy of their anti-piracy/anti-tampering technology which may be correlated to the technology used within your existing network.

Traditionally, this was done via human analysts to ensure proper tagging of the data for information ranging from video quality to broadcast information. Nowadays, this is performed via machine learning/image recognition which dramatically speeds up and improves the overall accuracy and quality of the piracy information provider service providers with an even more accurate picture of the piracy landscape.

4. Watermarking

Watermarking is the deliberate injection of information into a piece of content, with the intent of identifying its source upon later identification. Watermarking can occur with varying layers of granularity, with session based watermarking being the most useful. With each session identified with its own identifier, it is a simple forensic exercise to trace a piece of content back to the leaking device and mitigate accordingly.

5. OTT Credential Theft Monitoring

There are a variety of OTT credential theft mechanisms.

First is the sharing of OTT credentials between friends and family members. This takes a variety of forms from the relatively passive form of simply sharing user identification and password information to actively adding the devices of friends and family to one's active device list.

Second is the sharing of credentials for potential profit. For example, a university student may "rent" out OTT credentials to others to earn a little money on the side as they know their own parents/grandparents will not use any of the provided OTT credentials.

Third is the most nefarious form where hackers attack and obtain credentials via hacks of applications or via man-in-the-middle attacks via various public wireless hotspots. These fraudulently obtained credentials are, in turn, offered via dark web websites to members of the public.

All of these result in a major loss of potential subscribers to Operators in addition to churn which can be attributed to web streamed content re-broadcasts.

Addressing this specific mode of content loss requires a variety of tools. Restricting the number of concurrent streams per household via Rights Management concurrent stream control is one popular method. This is often supplemented by adding geo restrictions and controlling the number of registered devices per account to control access. To further protect against unplanned credential sharing, these measures are supplemented via application security such as cryptographic protection and communication as well as big data research to better track, visualize, and react to changes in OTT content usage to better react in real time to unwanted content theft.

6. Peer to Peer Monitoring and Analysis

P2P technology is typically used to share non-live content such as TV shows and films, but it can be used to also share live content via technologies such as Acestream and SOPCast. As a result, P2P technology applies to both nonlive and live content.

P2P non live content is typically shared via Torrent index sites. Bit Torrent is the dominant protocol in the P2P space so users search these Torrent index sites to obtain P2P infohashes. Once they have obtained this information, they use a Bit Torrent client to join a "swarm" and obtain all the necessary file parts from peers in their network.

P2P Live content is shared via SOPCast and Acestream. These addresses are shared via a variety of mechanisms from websites to social media. Social Media platforms such as Twitter and Reddit are fast becoming a preferred method for distributing this real time P2P piracy information.

To cover both forms of P2P piracy, crawlers for both the Torrent index sites as well as social media sites must be built. Using metadata analysis, the number of potential infringements must be quickly sifted so that an optimal number of swarms can be joined for video capture to enable video fingerprinting identification. This technology is key for enabling rapid cost effective anti piracy services for both real time and VoD.

In addition, P2P data can be collected to uniquely understand demand on an aggregate and geographic basis. This business intelligence information, in turn, can be used by operators to determine the potentially elasticity of piracy in their regions of interest in response to marketing, pricing, or product campaigns.

Mitigation

Mitigation of content can take many forms, including legal, based upon the nature of the detection. Upon the detection of a known piece of content, the issuance of a takedown notice is a common approach. In the case of real time content such as sporting events, these notices must be issued quickly otherwise the they have little effect.

On a more technical level, assuming techniques such as watermarking are used, individual devices can be targeted for disablement. If an operator can identify the source of a leak, cutting off the client is the most direct solution. This approach is somewhat risky as the client must be identified as a legitimate pirate point. Cutting off legitimate users from their service never ends well for the operator.

Conclusion

We've historically been living with piracy in parallel with our video delivery systems and making the value judgement of how much security is enough. The easy availability of tools and devices make piracy easier than effort and is feeding a monetized piracy industry that is mounting a real challenge to video operators moving forward.

In order to stem this hemorrhaging, we must now look beyond the device and identify our own content in the wild and take appropriate actions to mitigate its theft. Luckily, as discussed herein, the technologies and techniques exist and are being deployed on operator- wide scales.

Abbreviations

HDCP	high-bandwidth digital content protection
HDMI	high definition multimedia interface
IP	internet protocol
JTAG	Joint Test Action Group
OTT	over-the-top
PC	personal computer
STB	set-top box
TV	television
UHD	ultra-high-definition
VoD	video on-demand

Bibliography & References

Motion Pictures Laboratories, Inc. (2010, May). levels of verification for P2P scanning, version 2.0.1. San Francisco, CA: Author.