IPTV 불법복제콘텐츠 필터링 기술 개발

Development of the filtering technology of illegal IPTV contents

김 종 안*, 김 종 흠*, 김 진 한*, 진 영 민*

Jong-An, Kim, Jong-Heum Kim, Jin-Han Kim and Young-Min, Chin

Abstract

In this paper, both various content filtering methods and technological trends are briefly reviewed. A novel method is provided for preventing the illegal copies of copyright IPTV contents from being propagated through the P2P sites and web harddisk services.

Keywords : content filtering, forensic marking, fingerprint, watermarking

I. 서 론

국내 IPTV 서비스 가입자는 2009년 5월 현재 160만을 넘어섰다[1]. IPTV 서비스는 VOD 서비스와 실시간 TV 방송서비스로 구분되며, IPTV 사업자는 케이블TV 및 위성 방송사와의 우위경쟁력 확보와 수익 증대를 위해서 양질의 VOD 콘텐츠 수급에 많은 역량을 투입하고 있다. 이러한 예는 개봉관에서 상영을 마친 영화를 DVD로 출시되기 전에 IPTV VOD서비스로 제공하는 추세에서 찾아 볼 수 있다. 그러나 IPTV에서 서비스되는 신규 개봉영화콘텐츠가 불법 복제된 IPTV 버전 불법영화파일이 인터넷 파일공유사이트(P2P나 웹하드 사이트 등)를 통해 공유됨으로써 IPTV 서비스 사업자는 예상만큼 큰 효과를 거둘 수 없게 되었다. IPTV버전 영화파일은 IPTV 셋톱박스의 아날로그 출력단자(컴포넌트: Y, Pb, Pr, 컴포지트 등)에서 나오는 영상신호를 저렴한 HDTV 수신카드를 이용하여 디지털영상신호로 변환한 다음에, 인코딩하여 파일로 저장함으로써 손쉽게 생성될 수 있다.

이러한 IPTV 콘텐츠의 불법 복제 및 불법 유통을 추적을 위해 콘텐츠가 재생되는 시점에 디지털 워터마킹 기술을 이용하여 콘텐츠 사용자 정보를 삽입하고, 인터넷에서 불법 유통되는 디지털 콘텐츠로부터 삽입한 워터마킹 정보를 추출하여 콘텐츠를 불법 캡처 행위자를 알아내는 포렌식 마킹(Forensic Marking 이하 FM) 기술이 IPTV 서비스 최근 채택되고 있다.

FM 기술을 사용하면 불법복제후 유통되고 있는 IPTV 콘텐츠를 이용하여 불법복제자를 추적할 수는 있지만, 콘텐츠 유출로 인한 IPTV 사업자의 매출감소에는 큰 도움이 되지 못한다. FM 기술의 단점을 보완할 수 있고 인터넷 파일공유사이트에서의 불법복제물의 업로딩을 차단하기 위한 콘텐츠 필터링 기술에 대한 연구와 필터링 서비스 구축이 활발히 진행 중이다. 본 논문에서는 콘텐츠 필터링기술 종류와 국내외 기술 동향 등에 대하여 살펴보도록 하고, IPTV 불법복제콘텐츠의 불법공유를 효과적으로 차단할 수 있는 시스템을 제시하고자 한다.

접수일자 : 2009년 7월 18일
최종완료 : 2009년 8월 7일
*KT 중앙연구소
교신저자, E-mail : joankim@kt.com

II. 콘텐츠 필터링 기술

콘텐츠 필터링은 인터넷 파일공유사이트 등에서 콘텐츠의 저작권 침해 여부에 따라 파일전송을 제어하는데 사용하는 기술이다. 콘텐츠 필터링기술은 크게 콘텐츠 제목 등 문자열을 이용하는 텍스트 혹은 키워드(keyword) 기반 필터링, 파일마다 존재하는 해시(hash)값을 비교한 해시기반 필터링 그리고 콘텐츠의 특징정보를 이용한 특징점(feature) 기반 필터링 등으로 분류할 수 있다.

1. 필터링 기술 구분

가. 텍스트(키워드) 기반 필터링

텍스트 기반 필터링 기술은 콘텐츠 필터링작업(콘텐츠의 업로드나 다운로드를 허용하거나 불허하는 작업)에 콘텐츠 파일명이나 콘텐츠 제목명 등 주요단어(키워드)의 텍스트(문자열)를 이용하는 기술이다. 텍스트 기반 필터링은 사용자가 입력한 문자열과 저작물 제목 등으로 등록되어 있는 문자열들과의 비교를 통하여 간단히 저작물 침해 여부를 판별할 수 있어 처리 속도가 빠른 장점이 있으나 파일명을 바꾸면 필터링 작업을 우회할 수 있는 단점이 있다.

파일공유사이트에서 mp3 확장자 명을 가진 음악 콘텐츠는 공유가 금지 되지만, 음악콘텐츠의 파일명을 바꾸고, 확장자를 zip으로 바꿀 경우에 공유가 가능한 이유가 텍스트 필터링을 사용하기 때문이다[2-3].

나. 해시(hash) 기반 필터링

해시 기반 필터링은 해시함수를 이용하여 필터링 작업을 한다. 해시함수는 디지털 콘텐츠 파일을 입력하여 해시

값인 짧은 "전자지문(Signature)"를 만들어 낸다. 해시 함수로는 MD4, MD5, SHA 등을 사용하며, 좋은 해시 함수는 동일한 콘텐츠에 대해서는 같은 해시값을 출력하고, 서로 다른 콘텐츠에 대해 동일한 해시값을 출력하는 빈도수가 적은 함수이다.

해시 함수를 이용한 필터링은 저작권 보호 콘텐츠들에 대한 해시값을 데이터베이스에 저장한 다음, 공유하고자 하는 콘텐츠를 파일공유사이트에 업로드하는 시점에 서버 단에서 이 콘텐츠에 대한 해시값을 계산하여 데이터베이스의 값들과 비교하여 필터링을 수행하는 방식이다. 해시 기반 필터링은 텍스트 기반 필터링의 취약점인 제목 변경 등에 강인하고, 해시함수의 계산이 빨라 저작물 콘텐츠들에 대한 해시 데이터베이스 구축이 용이한 장점이 있다. 하지만, 해시값은 콘텐츠 파일의 미세한 변화에도 그 값이 달라지기 때문에 원본 저작물콘텐츠를 사용자가 압축포맷 변환, 단축 등 파일크기 변환 등에 상대적으로 취약하여 이들 변종파일에 대한 해시 데이터베이스를 지속적으로 현행화해야만 하는 단점이 있다.

다. 특징점(feature) 기반 필터링

특징점 기반 필터링 혹은 내용기반 필터링은 콘텐츠 데이터 고유의 특징을 추출하여 이를 기반으로 파일 필터링을 하는 기법이다. 특징점(feature)은 내용(content), DNA, 지문(fingerprint) 등 다양한 용어로 불린다. 콘텐츠의 특징 추출에는 컬러 정보, 명암정보, 모션정보, 주파수대역 에너지 값 등 다양한 정보를 사용한다.[3][4]

특징점 기반 필터링은 저작물 콘텐츠의 고유한 특징을 필터링에 사용하므로, 포맷, 파일명 변경 등에도 강인한 장점이 있지만, 해시기반 필터링과 같이 저작물에 대한 특징점 DB(Database)를 구축해야 한다. 또한, 특징점 추출은 앞서 언급한 텍스트나 해시값의 경우보다 계산량이 많고, 특징점 저장에 할당되는 용량도 많이 드는 단점이 있다.

2. 콘텐츠 필터링 시스템

앞서 언급한 텍스트, 해시, 특징점 기반 콘텐츠 필터링 시스템의 구조는 그림 1과 같다. 콘텐츠 필터링 시스템은 저작권침해 판별시스템, 웹하드서비스 서버, 그리고 웹하드서비스 사용자로 크게 구분된다.

저작권침해 판별시스템은 저작권보호를 받고자 하는 콘텐츠를 이용하여 콘텐츠 제목, 해시값, 특징점 등을 추출하는 콘텐츠 대표정보 추출프로그램, 그리고 이러한 추출된 정보로 구축된 데이터베이스(DB), 웹하드서비스 업체에서 전송해온 콘텐츠 대표정보값을 이용하여 DB를 검색하여 저작권침해여부를 판별하는 프로그램으로 이루어진다.

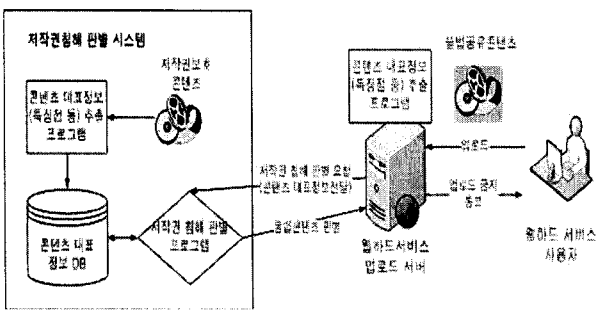


그림 1. 콘텐츠 필터링 시스템 구조

웹하드서비스 사용자는 자신이 확보한 불법공유콘텐츠의 업로드를 시도하면, 웹하드서비스업체의 업로드서버에서는 저작권침해 판별시스템의 것과 동일한 콘텐츠 대표정보값 추출프로그램을 이용하여 업로드 시도 콘텐츠에 대한 콘텐츠 정보값을 추출한다. 이렇게 추출된 정보값을 저작권침해 판별시스템의 저작권침해 판별 프로그램에 전송하면서 저작권 침해 여부를 문의한다. 저작권을 침해하는 불법콘텐츠라는 결과를 통보받으면 웹하드 업로드서버는 해당 콘텐츠의 업로드를 진행할 수 없다는 경고메시지를 웹하드서비스 사용자에게 전달한다. 결과적으로 웹하드 서비스 사용자는 저작권보호를 받고 있는 콘텐츠를 공유할 수 없게 된다.

저작권침해 판별시스템은 저작권보호센터와 같은 공공성을 지닌 국가기관에서 관리 감독하에 운영하는 것이 타당하다.

3. 콘텐츠 필터링 기술 현황

저작권법 제104조에 따르면 "특수한 유형의 온라인서비스제공자(OSP, 예를 들어 웹하드서비스 사업자) "는 권리자의 요청이 있는 경우 해당 저작물 등의 불법적인 전송을 차단하는 기술적인 조치를 취해야 한다고 명시되어 있다. 저작권법 상에서 기술조치에 대해 특정기술을 명시하지는 않았지만, 현재 대부분의 특수 유형의 OSP(Online Service Provider)는 금칙어 설정을 통한 검색제한 방식 즉 텍스트 기반 필터링을 적용하고 있다.

OSP에서는 중복 파일 제거를 위해 해시를 사용하고 있다. 하지만, 동일저작물이라도 수십 개의 불법 파일이 존재하여 실질적으로 해시값만으로 필터링을 하기에는 무리가 있다. 특징점 DB를 이용한 콘텐츠 필터링은 모든 필터링 방식에서 가장 강력한 필터링 방식이나 특징점 DB 구축을 위한 저작물 확보가 어렵고 운영비용이 부담스러운 상태이다.

표 1은 국내 콘텐츠 필터링 솔루션 현황과 표 2는 해외 콘텐츠 필터링 솔루션 현황을 각각 나타낸다[2,5].

표 1. 국내 콘텐츠 필터링 기술업체

솔루션업체	솔루션 설명
ETRI	- 음원/영상인식원천 기술 개발 - 클루넷 등 기술이전
CK&B	- 방송 프로그램의 불법 복제 모니터링 및 추적 시스템 개발 - 워터마킹, 로고인식, DNA 기반
마크애니	- 동영상, 음원 DNA 필터링 기술 보유 - 엠넷미디어(뮤직포탈), 씨네로닷컴(온라인 영화)에 적용
뮤레카	- 특징점 기반 오디오/비디오 검색 기술 - 싸이월드의 음원저작물에 대한 필터링 서비스 제공 - 위디스크에 불법콘텐츠 필터링 시스템 제공 - 동영상 필터링 기술 무료 제공 선언

솔루션업체	솔루션 설명
비티워스	- 스크립트 기반 웹검색 기술을 이용한 불법저작물 검색
엔써즈	- 오디오/비디오 핑거프린팅 기술 보유 - 네이버와 다음 음원 필터링 서비스에 적용, 씬지오에 음원 및 동영상 필터링서비스제공
위디랩	- 동영상 장면의 변화량을 특징점으로 이용한 "프리즈마(PRISMA)" 검색엔진 출시
지란지교소프트	- 유해영상차단 솔루션(해시 기반)
클루넷	- 오디오DNA필터링 서비스 - KTH, 짱파일, 클럽포스, 엠파일, 피디팝, 온디스크 등 22개사와 계약체결

표 2. 해외 콘텐츠 필터링 기술업체

솔루션업체	솔루션 설명
AudibleMagic	- 미국 소니사의 자회사로 YouTube에 콘텐츠 식별서비스 제공 - Myspace.com의 콘텐츠 식별서비스 제공
DtecNet	- ISP에서 적용할 수 있는 불법콘텐츠 필터링 기술 보유, 네트워크 트레픽을 검사하여 불법콘텐츠 차단
Gracenote	- 오디오/비디오 인식기술 보유
Phillips	- 오디오/비디오 인식기술 보유 - "MEDIAHEDGE" 서비스 개발 - CIvolution이란 자회사에서 기술 개발
소니-에릭슨	- 모바일 사용자를 위한 TrackID 서비스
Vobile	- "VidoeDNA" 검색 기술 보유 - "VideoTracker" - AT&T가 네트워크 동영상 필터링 서비스 제공을 위해 자본 투자(2007.11)

III. IPTV 불법복제콘텐츠 필터링 시스템

앞서 언급했듯이 IPTV 사업자는 신규 수익원으로 개봉관에서 상영을 마친 영화를 DVD 출시 전까지 IPTV에서 서비스하는 새로운 콘텐츠 비즈니스 사슬(chain)을 생성하였으나, IPTV 셋톱박스의 아날로그 출력신호에 대한 콘텐츠 보호솔루션 미비로 IPTV 버전의 불법콘텐츠들이 인터넷파일공유사이트에서 유포됨에 따라, 기존의 DVD 시장마저 위협을 받게 되었다. 이러한 사실은 현존하는 디지털콘텐츠 필터링 기술(텍스트기반, 해시기반, 특징점 기반 필터링 등)의 한계점을 보여준다. 특징점 기반 필터링은 효과가 높은 것은 사실이나 특징점 추출프로그램이 무수히 많은 인터넷 공유파일사이트의 서버에 설치되고 특징점 DB 서버와 연동되어야 효과가 있는 솔루션이고 특징점 기반 업체 마다 독자적인 특징점 추출솔루션을 보유하다 보니 특징점 DB의 종류도 업체 수만큼 많다고 볼 수 있다. 저작권법 제104조에 의거하여 웹하드디스크 업체들이 특징점 기반 필터링 솔루션을 설치한다 하더라도 각 업체들의 특징점 DB를 서로 공유하지 않는 한 특징점을 이용한 불법콘텐츠 필터링은 큰 효과를 볼 수 없다.

본 논문에서 이러한 특징점기반 필터링솔루션의 단점을 보완할 수 있는 워터마크 기반 IPTV 불법복제콘텐츠 필터링 시스템(그림 2 참조)을 제안하고자 한다.

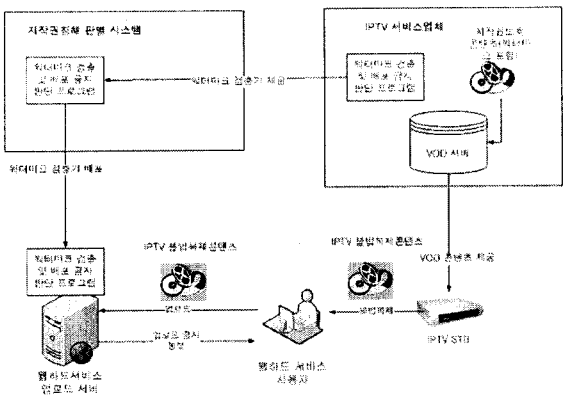


그림 2. IPTV 불법복제콘텐츠 필터링 시스템

워터마크기반 IPTV 콘텐츠 필터링 시스템에서 IPTV 서비스업체는 IPTV VOD콘텐츠에 저작권 관련 워터마크를 삽입한 후 IPTV STB에 제공하므로, 해당 콘텐츠가 불법 복제되어도 워터마크 정보가 들어있게 된다. IPTV 서비스업체는 이렇게 삽입된 워터마크를 검출하고 저작권관련 판단을 할 수 있는 프로그램을 저작권 침해관리 시스템에 배포요청을 하면, 인터넷파일공유사이트 업체가 이 프로그램을 배포 받아 설치한다. 웹하드사용자가 업로드하는 IPTV 불법콘텐츠에 대해 상기 프로그램을 이용하여 저작권 정보를 검출한 다음 업로드 허락여부를 결정한다.

이러한 워터마크 기반 필터링은 특징점기반 필터링에서 필요한 DB 구축 및 연동이 필요 없는 장점이 있으나, 워터마크 검출에 시간이 많이 소요되어 웹하드서비스 사용자에게 불편을 초래할 수 있다.

IV. 결 론

본 논문에서는 IPTV 콘텐츠의 복제 및 유출 원인에 대해서 살펴보았고, 불법복제 콘텐츠를 인터넷파일 공유사이트에서 차단할 수 있는 콘텐츠 필터링 기술과 국내외 기술동향에 대해서 알아보았다. 또한, IPTV 불법복제콘텐츠를 효과적으로 필터링 할 수 있는 시스템을 제시하였다.

[참 고 문 헌]

[1] KT 경제경영연구소, 2009년 7월 유무선 통신서비스 주요지표

[2] 한국정보통신기술협회(TTA), "콘텐츠 필터링 서비스 인터페이스," TTAK.KO-12.0070(2008-12-19)

[3] 김주섭, 남제호, "불법 복제 콘텐츠 필터링 기술 동향 분석," 한국방송공학회지, vol. 12, no. 4, pp. 55-63, 2007.

[4] 김주섭, 남제호, "DTV 방송프로그램의 온라인 불법전송 차단을 위한 비디오 복사본 검출 알고리즘," 방송공학회논문지, vol. 13, no. 5, pp. 662-676, 2008.

[5] 이성환, "불법 복제물의 필터링 기술 동향", TTA DRM 프로젝트그룹(PG506) 2009PG506-082, 2009.5.

김 종 안

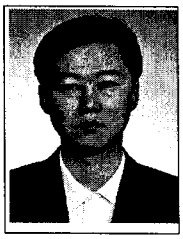
1984년 고려대학교 전기공학과 졸업
1988년 고려대학교 전자공학과(공학석사)
1997년 미국미네소타대학교 컴퓨터공학과(박사 수료)
1988년~현재 KT 중앙연구소 부장
<관심분야> DRM/CAS, Forensic Marking, IPTV 콘텐츠 보호
<e-mail> joankim@kt.com

김 종 흠

1999년 포항공과대학교 전자전기공학과 졸업
2001년 포항공과대학교 전기전자공학과(공학석사)
2004년~현재 KT 중앙연구소 사원
<관심분야> DRM, FM, CAS, 암호화 알고리즘
<e-mail> jinah0719@kt.com

김 진 한

1986년 고려대학교 전자공학과 졸업
1988년 한국과학기술원 전기및전자공학과(공학석사)
1992년 한국과학기술원 전기및전자공(공학박사)
1992년~현재 KT 중앙연구소 부장
<관심분야> DRM, IPTV, 미디어 공통플랫폼, On-Demand 서비스, 디지털컨텐츠, WiBro
<e-mail> jinhan@kt.com



진 영 민

1983년 연세대학교 학사
1985년 연세대학교 석사
1989년 연세대학교 공학박사
1990년~현재 KT 중앙연구소 상무보
<관심분야> IPTV, 플랫폼/STB, CAS/DRM, 양방향서비스 기술, UX
<e-mail> ymchin@kt.com