

디지털콘텐츠 고유 식별자를 이용한 불법콘텐츠 추적연구

강상욱^a, 박승범^b, 강경훈^c, 강호갑^d, 이규정^e

^a 한국정보사회진흥원 책임연구원
100-170, 서울시 중구 무교동 77번지 NIA빌딩 u비즈니스탑
Tel: +82-2-2131-0443, Fax: +82-2-2131-0303, E-mail: sukang@nia.or.kr

^b 한국정보사회진흥원 선임연구원
100-170, 서울시 중구 무교동 77번지 NIA빌딩 u비즈니스탑
Tel: +82-2-2131-0449, Fax: +82-2-2131-0309, E-mail: parksb@nia.or.kr

^c 한국정보사회진흥원 선임연구원
100-170, 서울시 중구 무교동 77번지 NIA빌딩 u비즈니스탑
Tel: +82-2-2131-0484, Fax: +82-2-2131-0309, E-mail: emptyee@nia.or.kr

^d DRM Inside 연구소장
100-170, 서울특별시 송파구 가락본동 98 두산위브센터 410호
Tel: +82-2-2131-0141, Fax: +82-2-2131-0309, E-mail: hgkang@drminside.com

^e 한국정보사회진흥원 e비즈니스탑 팀장
100-170, 서울시 중구 무교동 77번지 NIA빌딩 u비즈니스탑
Tel: +82-2-2131-0141, Fax: +82-2-2131-0309, E-mail: lkj@nia.or.kr

Abstract

본 연구는 기술적인 측면에서 불법 콘텐츠 유통을 방지하기 위한 현행 추적기술들을 조사해 보고, 이들이 갖고 있는 문제점을 살펴본 후 디지털콘텐츠 식별체계 국가표준으로 채택된 한국정보사회진흥원의 UCI(Universal/Ubiquitous Content Identifier) 시스템과의 연계를 통해 문제점을 해결할 수 있는 방안을 연구하여 불법콘텐츠 사용 방지를 위한 기술적, 정책적 방안을 도출하며, 결론으로서, 이러한 불법콘텐츠 추적기술들과 UCI와의 연계를 통해 추적시스템에서의 CP 정보 중복 저장 방지 및 관리의 편리성을 향상시킬 수 있는 방안을 제시할 것이다.

Keywords: 식별자, 불법콘텐츠 추적, UCI

1. 서론

디지털콘텐츠 유통이 활발해지고 디지털콘텐츠의 복제 및 재가공이 쉬워짐에 따라 콘텐츠의 불법복제로 인한 저작권 침해 문제가 사회적 문제점으로 대두되고 있다. 이에 따라 콘텐츠 제공자들은 콘텐츠의 불법복제 및 부정 이용을 막기 위해 DRM, CAS 등 다양한 저작권 보호기술을 사용하는 등 적극적 콘텐츠 보호 활동을 하고 있다. 그러나 이러한 노력에도 불구하고 수많은 콘텐츠들이 불법으로 제작/유통되고 있어 정상적으로 유통되는 많은 콘텐츠들이 자리를 잡지 못하고 있다. 정상적인 콘텐츠보다 불법콘텐츠가

더욱 쉽게 전파되고 기승을 부리는 이유는 다음과 같이 정리해 볼 수 있다.

첫째, 디지털콘텐츠에 대한 복제 및 재배포 용이성을 들 수 있다. 현재의 컴퓨터 시스템에서는 디지털콘텐츠에 대한 복제 및 재배포가 매우 간단하다. 따라서 누구나 인터넷에 연결된 PC만 있으면 자신이 가지고 있는 디지털콘텐츠를 복제해서 인터넷을 통해 배포하는 것이 가능하고 유통되는 콘텐츠가 별도의 배포경로에 대한 정보를 가지고 있지 않기 때문에 결과적으로 불법 배포자들의 익명성을 제공해주고 있다.

둘째, 불법콘텐츠 추적의 어려움을 들 수 있다. 콘텐츠가 불법적으로 유통되었을 경우에는 대부분 최종 배포자 또는 사용자만을 확인할 수 있다. 최초의 불법복제 및 배포의 행위를 수행한 자나 중간 유통자들을 시스템적으로 알아내기는 쉽지 않다. 최근 들어 핑거프린팅 기술을 이용하여 최초의 불법복제자를 검출하는 시스템을 사용하고는 있지만, 이 기술도 중간 유통자들을 알아내지는 못한다.

마지막으로, 양질의 유료 콘텐츠 부재 및 합법적 콘텐츠 사용에 대한 의식의 미약을 원인으로 제시할 수 있는데, 대다수 사용자들은 동일한 콘텐츠가 합법적인 사이트에도 존재하고 불법적인 유통 사이트에도 존재를 한다면 어느 쪽을 선택할지 고민하지 않을 수 없다. 그러나 원하는 콘텐츠가

불법적인 유통 사이트에서만 존재한다면 선택의 여지가 없을 것이다. 특히 영화 콘텐츠의 경우 서비스를 하기 위해서는 배급사와 같은 저작권자와의 협의가 필요하고 모든 영화에 대해 저작권을 확보하기가 쉽지 않기 때문에 유료 영화 VOD 서비스 사이트는 양질의 콘텐츠에 대한 확보 부담을 안고 있다. 따라서 일부 사용자들은 자신이 보고 싶어 하는 많은 영화를 P2P 사이트를 통해서 불법적으로 다운로드 받고 있는 현상이 나타나고 있다.

상기 언급한 바와 같이 불법콘텐츠 유통이 쉽게 근절되지 않고 있는 상황에서 이를 추적하여 법적인 조치를 취하기 위한 불법 콘텐츠 추적기술들이 활발히 연구되고 있다.

본 연구는 기술적인 측면에서 불법 콘텐츠 유통을 방지하기 위한 현행 추적기술들을 조사해 보고, 이들이 갖고 있는 문제점을 살펴본 후 상용 디지털콘텐츠 식별시스템과의 연계를 통해 문제점을 해결할 수 있는 방안에 대해 논하고자 한다

2. 현행 불법 콘텐츠 추적기술의 문제점

2.1 콘텐츠 불법복제 및 유통 형태

불법 콘텐츠는 디지털콘텐츠의 불법적인 유통으로부터 발생된다. 개인적인 사용을 목적으로 하는 사적복제는 라이선스가 있는 콘텐츠에 대해서 법률적으로도 허용을 하고 있지만 사적복제된 콘텐츠가 개인적인 용도 이외의 목적으로 유통된다면 이는 별도의 라이선스를 필요로 하는 것이기 때문에 복제와 유통의 행위가 모두 불법이 된다. 또한 라이선스가 없는 콘텐츠는 사적인 용도를 위한 복제라 하더라도 사실상 모두 불법복제에 해당한다.

불법콘텐츠는 디지털콘텐츠를 인터넷에 무단 공유하고 이를 다운로드 받는 과정을 통해 불법 유통되고 있으며, 공유 방법으로는 주로 P2P 사이트가 많이 남용되고 있다. 불법으로 복제된 콘텐츠는 개인적인 용도로 사용되기도 하고 유사한 경로를 통해 다시 공유되기도 하며, 경우에 따라서는 자신의 혹은 회사의 홈페이지를 꾸미기 위한 공적인 용도로 도용되기도 한다.

사용자의 불법 복제 및 유통 행위는 DRM 등의 콘텐츠 보호기술을 통하여 기술적으로 어렵게 만들어 놓을 수는 있다. 그러나 아무리 보호 기술이 강력하다고 하더라도 아날로그 홀(Analog Hole)을 포함한 여러 가지 문제점들로 인하여 100% 방지하는 것은 불가능한 것으로 알려지고 있다. 따라서 디지털콘텐츠에 대한 불법복제 및 유통은 기술적인 보호조치를 포함해서 사용자들의 불법콘텐츠 사용에 대한 인식변화를 위해

법적/제도적인 방지대책이 뒷받침 되어야 하며, 법적 근거가 마련되어 있는 경우 집행 차원의 불법콘텐츠 추적에 대한 시스템 도입도 필요하다고 볼 수 있다.

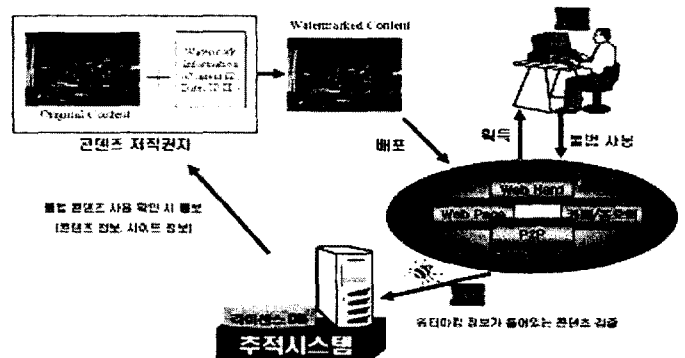
2.2 현행 불법콘텐츠 추적 기술

현재까지 알려진 불법콘텐츠 추적 기술에는 다음과 같은 것들이 있다.

2.2.1 상용 콘텐츠에 대한 추적기술

2.2.1.1 워터마킹 기법을 이용한 추적기술

워터마킹 기법을 이용한 추적기술은 콘텐츠 배포시 내부에 저작권자에 대한 정보를 워터마킹 기술을 이용해 삽입하여 향후 저작권 분쟁이 발생하였을 경우 증거 자료로 삼을 수 있는 기술이다. 예를 들어 한 방송사에서 제작된 동영상 자료의 내부에 방송사의 정보를 워터마킹 형태로 삽입해 놓으면 타 방송사가 이를 무단 도용하여 또 다른 방송물을 제작하였을 경우 해당 동영상의 워터마킹 정보를 추출하여 도용의 증거로 제시할 수 있는 추적방식이다. 또한 이 방식은 Web Spider¹와 같은 불법 콘텐츠 추적시스템 통해 자사의 워터마킹 정보가 들어있는 콘텐츠가 인터넷상에서 발견되었을 경우 라이선스 발급현황을 조사하여 불법 사용여부를 판단하는 형태로도 사용되고 있다. 그러나 이 방식은 최종적으로 불법 도용한 사용자가 누구인지를 판별할 때 사용되는 추적방식으로 최초 배포자를 가려내는 데는 활용될 수 없는 단점을 가지고 있다. 따라서 배포에 대한 제한을 두지는 않지만 공공의 목적으로 사용하거나 인터넷에 게시를 하고자 할 경우 저작권자의 동의를 필요로 하는 콘텐츠에 주로 활용되고 있는 추적 기술이다.



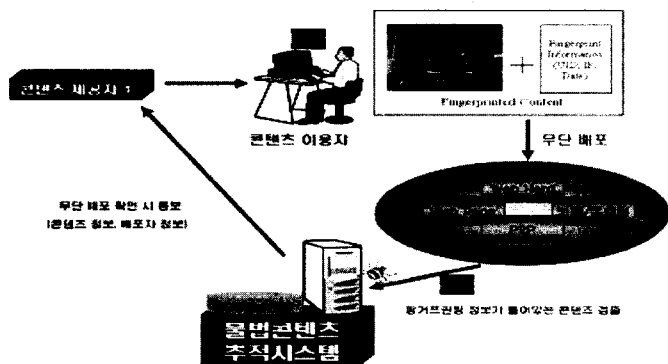
(그림 2-1) 워터마킹 기법을 이용한 추적기술

2.2.1.2 핑거프린팅 기법을 이용한 추적기술

핑거프린팅 기법을 이용한 추적기술은 사용자가

¹ Web Spider: 한국전자통신연구원에서 개발한 불법콘텐츠 추적 시스템..

특정 콘텐츠를 이용할 때 렌더링 또는 출력되는 콘텐츠의 내부에 사용자에게 대한 정보를 핑거프린팅 형태로 삽입함으로써 만일 무단 배포 행위가 발생하였을 경우 최초 불법복제 및 배포 행위를 수행한 자에 대한 정보를 획득하기 위한 기술이다. 예를 들면, 온라인 영화 서비스 사업자로부터 구매를 통해 다운로드 된 영화 콘텐츠를 사용자가 이용하고자 하는 시점에 사용자에게 대한 정보를 핑거프린팅 기술을 이용해 삽입해 놓으면 사용자가 이를 무단 배포하였을 경우 콘텐츠 내부에 삽입된 핑거프린팅 정보를 추출하여 누가 최초 배포자인지 추적할 수 있는 방식이다. 이 방식은 주로 Web Spider 같은 추적시스템을 통해 핑거프린팅 정보가 들어있는 콘텐츠가 인터넷상에서 발견되었을 경우 라이선스 발급현황을 조사하여, 불법 사용 및 배포 여부를 판단하는 형태로도 사용되고 있다. 이 방식은 최종적으로 불법 도용한 사용자의 판별뿐 아니라 최초 배포자를 가려내는 데에도 유용하게 사용되는 추적기술이다.



(그림 2-2) 핑거프린팅 기법을 이용한 추적기술

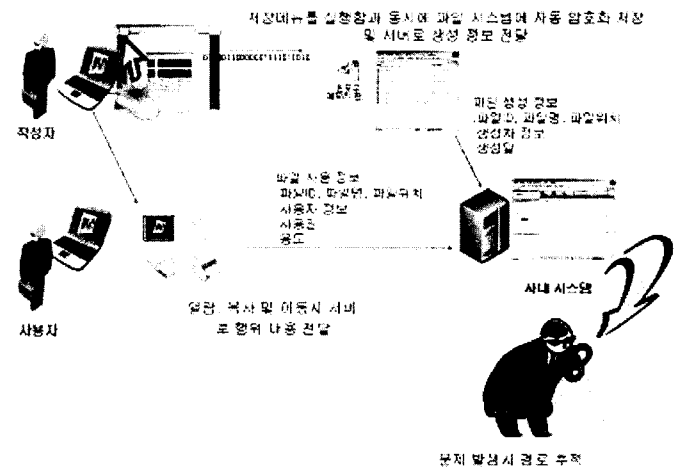
2.2.2 기업내 비 인가된 문서 유통 추적기술

기업들은 디지털화되어 있는 기업 내부의 중요 문서나 도면이 외부로 유출되었을 경우 기업의 이미지뿐만 아니라 막대한 경제적 손실이 발생할 수 있기 때문에 기밀문서의 무단 배포를 철저히 통제하고 있으며 비 인가된 문서에 대한 유통 경로를 추적하기 위해 다양한 PC 보안 기술을 도입하고 있다.

이 기술은 사용자가 문서나 도면을 작성하거나 복사를 하게 될 경우, 컴퓨터상의 파일 시스템에서 자동으로 사용자 정보와 파일 정보, 그리고 파일의 저장위치 정보를 서버로 전달하여 향후 문제가 발생하였을 경우 유통의 경로를 포괄적으로 추적해 주는 기술이다. 이 기술은 특정 문서에 대한 최초 작성자부터 중간 유통자 및 최종 유포자 까지 해당 문서의 모든 이동 경로를 추적할 수 있다는 장점을 가지고 있다.

그러나 사용자 시스템에는 반드시 PC 보안 제품이 설치되어야 하고 통제를 벗어난 개인적인 컴퓨터의

사용을 방지하기 위해 주로 PC-DRM 기술과 접목되어야 하기 때문에 특정 기업내부에서 사용될 수는 있어도 상업적인 용도로 사용되기에는 무리가 있는 기술이다.



(그림 2-3) 기업내 비 인가된 문서 유통 추적기술

2.2.3 현행 불법콘텐츠 추적 기술의 문제점

현재 사용되고 있는 불법콘텐츠 추적기술은 각각 장단점을 가지고 있는데, 전반적으로 '위터마팅 기술' - '핑거프린팅 기술' - '파일시스템 통제 기술' 순으로 추적의 범위가 넓어지는 장점이 있는 반면 시스템 부하 측면에서는 점점 부담을 주고 있다. 특히 기업내 비인가 문서 추적시스템에 활용되고 있는 파일시스템 통제 기술은 파일 시스템에 별도의 제어기술이 추가되어야 하기 때문에 사용자 PC에 상당한 부담을 주는 것으로 알려져 있다. 또한 상용 콘텐츠에 적용되는 불법 콘텐츠 추적 기술들은 워터마킹 또는 핑거프린팅정보에 삽입되는 콘텐츠식별자가 서비스업체별로 다르기 때문에 추적시스템이 별도의 부가정보를 저장하고 있어야 하기 때문에 정보의 중복저장과 정보 수정시 이중 작업이 필요하다는 단점을 가지고 있다. 다음은 현행 불법 콘텐츠추적 기술에 대한 장단점을 분석한 표이다.

<표 2-1> 현행 불법콘텐츠 추적 기술의 장단점

불법콘텐츠 추적에 사용되는 기술		장점	단점
상용 콘텐츠	위터마킹	한번의 워터마킹 삽입으로 지속적인 추적 관리 가능	최초 배포자, 중간 배포자를 추적할 수 없음 콘텐츠식별자 포맷이 서비스업체별로 다르기 때문에 추적시스템이 콘텐츠 식별자와

	핑거프린팅	최초 배포자 및 최종 사용자에 대한 정보를 추적하기에 용이	특정 사용자에게 배포시 마다 핑거프린팅 삽입 필요	관련된 콘텐츠 제공자를 찾기 위해서는 별도의 부가정보를 저장하고 있어야 함.
기업내 문서	파일시스템 통제	콘텐츠의 모든 유통경로 정보를 추적할 수 있음	- 파일 시스템 통제에 따른 시스템 부하 과중 - 통제를 벗어난 컴퓨터상 에서의 유통시 추적의 어려움 존재	

3. 불법 콘텐츠 추적기술과 디지털콘텐츠 식별체계 간 연계 방안

앞에서 살펴본 바와 같이 현행 불법 콘텐츠 추적기술은 기존의 콘텐츠 식별체계가 통일화되어 있지 않음으로 인해 추적시스템이 개별 식별체계와 이를 관리하는 콘텐츠 제공자에 대한 DB를 유지하고 있어야 되는 번거로움이 존재한다. 따라서 모든 콘텐츠들이 디지털콘텐츠 식별체계를 사용할 경우 추적시스템은 별도의 정보 DB를 관리하지 않고도 추출된 콘텐츠식별자와 디지털콘텐츠 식별 시스템을 통해 콘텐츠의 제공자가 누구인지를 알 수 있다. 디지털콘텐츠 식별체계란 온라인상에서 디지털콘텐츠의 효율적인 유통과 활용을 위해 콘텐츠 하나하나에 유일한 코드를 부여하고 이를 관리해 주는 체계로서, 본 연구에서는 디지털콘텐츠 식별체계 국가표준으로 채택된 한국정보사회진흥원의 UCI(Universal/Ubiquitous Content Identifier)² 시스템을 이용한 불법콘텐츠 추적시스템과의 연계를 살펴보기로 한다.

3.1 워터마킹 및 핑거프린팅 기술과의 연계

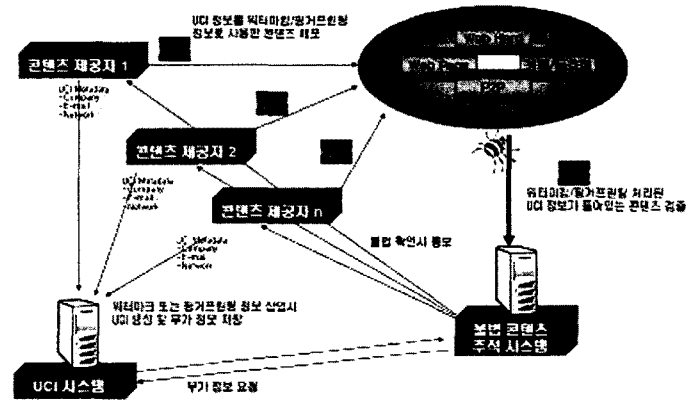
워터마킹 또는 핑거프린팅 기술을 이용한 현행 불법콘텐츠 추적시스템은 추적시스템을 기반으로 콘텐츠가 제작된 것이 아니기 때문에 추적의 대상이 되는 디지털콘텐츠들에 대한 식별체계가 통일되어 있지 않고, 이를 관리하는 곳도 각 서비스 사업자마다 다르기 때문에 추적시스템 자체에 식별자별로 이와 연관된 부가정보를 저장하고

2 UCI(Universal/Ubiquitous Content Identifier)는 URN 기반의 디지털 콘텐츠 식별체계로서 각 기관별·기업별 콘텐츠 식별체계에 상호호환성을 부여하는 최상위 국가표준 식별체계

※ URN(Uniform Resource Name): 디지털콘텐츠 자체에 상품바코드와 같은 영구불변의 식별코드를 부착하여 온라인상의 콘텐츠 유통에 활용하는 체계

관리해야 할 필요가 있다. 예를 들어 Web Spider에 의해 검출된 워터마킹 정보의 콘텐츠 아이디가 CompanyA::Content1 일 경우 이 콘텐츠가 합법적으로 사용되고 있는지 여부를 알기 위해서는 CompanyA에 Content1에 대한 라이선스가 콘텐츠가 게시된 곳에 발급이 되었는지 여부를 확인할 필요가 있다. 이때 추적시스템은 CompanyA의 정보시스템과 연동을 위한 부가정보(이메일 주소, 네트워크 주소 등)를 가지고 있어야 하며, 이 정보가 변경이 되는 경우에는 추적시스템의 DB 내용을 수정해야 되는 단점이 존재한다.

반면 UCI를 사용하는 콘텐츠의 경우에는 워터마킹 및 핑거프린팅 정보에 별도의 제공자 정보를 저장할 필요가 없으며, 검출된 정보로부터 UCI 시스템을 통해 제공자의 부가정보를 확인할 수 있기 때문에 정보의 중복 저장을 피할 수 있다. 또한 콘텐츠 제공자의 정보가 변경되었다 하더라도 추적시스템에는 아무런 영향을 주지 않고, 단지 UCI 시스템에서의 부가정보 변경만으로 시스템 운영이 가능하게 된다. 다음 그림은 UCI 시스템과 불법 콘텐츠 추적 시스템의 연동 모습을 보여준다.

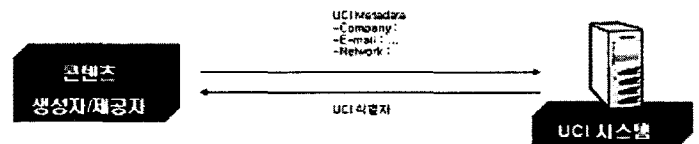


(그림 3-1) 불법콘텐츠 추적기술과 UCI 연계방안

UCI와 불법콘텐츠 추적시스템의 단계별 연계 방법은 아래와 같다.

3.2 콘텐츠 생성단계에서의 UCI 역할

콘텐츠를 생성할 때 콘텐츠에 대한 식별자를 UCI 시스템으로부터 받고, 콘텐츠에 대한 부가정보를 UCI 시스템에 등록한다. UCI 시스템으로부터 부여받은 식별자는 워터마킹이나 핑거프린팅 대상이 되는 콘텐츠의 식별자로 사용된다.



(그림 3-2) 콘텐츠 생성단계에서의 UCI 역할

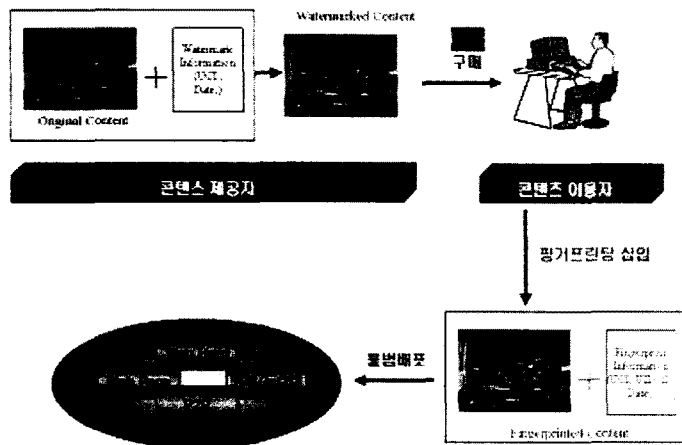
3.3 콘텐츠에 워터마킹이나 핑거프린팅 삽입단계에서의 UCI 역할

워터마킹이나 핑거프린팅시 삽입되는 콘텐츠 식별자는 UCI 식별자가 사용되며 제공업체에 대한 정보는 UCI 시스템이 관리하기 때문에 별도로 워터마킹이나 핑거프린팅 정보에 삽입하지 않는다.

<표 3-1> UCI와 연계된 불법콘텐츠 추적시스템에서의 삽입정

삽입되는 정보	워터마킹	핑거프린팅
현행 불법콘텐츠 추적시스템	<ul style="list-style-type: none"> •식별자 : DC001 •제공자 : CP_A •일자 : 2006.12.20 	<ul style="list-style-type: none"> •식별자 : DC001 •제공자 : CP_A •사용자 ID : User12734 •일자 : 2006.12.20
UCI와 연계된 불법콘텐츠 추적시스템	<ul style="list-style-type: none"> •식별자 : A999:CP_A:DC001 •일자 : 2006.12.20 	<ul style="list-style-type: none"> •식별자 : A999:CP_A:DC001 •사용자 ID : User12734 •일자 : 2006.12.20

다음 그림은 워터마킹 또는 핑거프린팅시에 삽입되는 UCI 정보와 UCI 시스템에 등록되는 부가정보의 예를 보여준다.

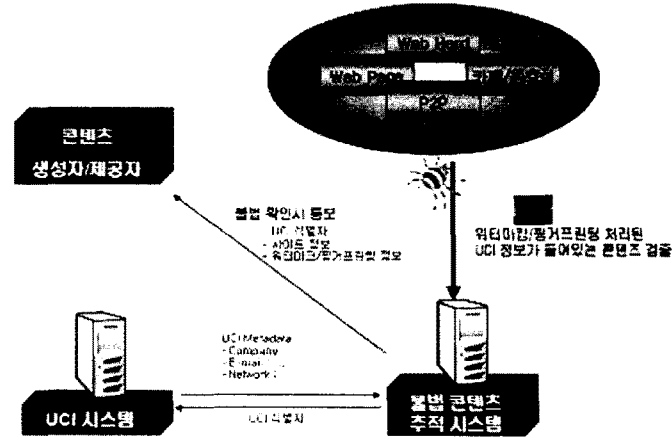


(그림 3-3) 워터마킹이나 핑거프린팅 삽입단계에서의 UCI 역할

3.4 추적시스템에 의해 검출된 정보 확인 단계의 UCI 역할

인터넷을 통해 검출된 정보에서 추적시스템은 UCI 식별자를 확인할 수 있으며, 이 정보를 UCI 시스템에 의뢰하면 이 콘텐츠에 대한 제공자 정보와 이 제공자에 대한 URL 정보를 얻을 수 있다. 이 정보를 바탕으로 추출된 워터마킹 정보 또는 핑거프린팅 정보를 콘텐츠 제공자에게 전달해 준다.

추적 시스템이 불법여부까지 확인하는 기능을 가진 경우에는 라이선스 발급 여부를 콘텐츠 제공업체에 문의한 후 불법이라 판단되었을 경우 위법사실을 통보해 줄 수 있다. 이때 콘텐츠 제공업체에서 라이선스 발급여부에 확인해 주는 시스템을 운영하고 있을 경우, 이에 대한 통신정보도 UCI 시스템을 통해 얻을 수 있다.



(그림 3-4) 검출된 정보 확인 단계에서의 UCI 역할

3.5 파일 시스템 통제 기술과 UCI 연계방안

주로 기업내 기밀문서에 대한 유통경로 추적에 사용되는 파일 시스템 통제 기술은 외부와의 연결을 원하지 않기 때문에 UCI 시스템과의 연동은 쉽지 않을 것으로 보인다.

4. 활용 방안

현재 한국정보통신표준협회의 표준화 기술위원회에서는 DRM Project Group(PG110)을 운영하고 있는데, 이곳에서 2006년부터 DRM과 핑거프린팅 인터페이스에 대한 표준화 과제를 진행하고 있다. 이 과제는 핑거프린팅과 DRM기술을 연동하여 DRM에 의한 보호 장치가 해체된 경우 불법으로 유포되는 콘텐츠에 대한 추적시스템 표준인터페이스 정의를 목적으로 하고 있다. 이 시스템에서 핑거프린팅 삽입시 들어가는 정보에는 콘텐츠 ID, CP 정보, 사용자 정보, 일자 등으로 본 연구에서 지적했던 CP 정보에 대한 중복 저장관리의 문제점을 가지고 있다. 따라서 UCI와의 연계를 통해 추적시스템에서의 CP 정보 중복 저장 방지 및 관리의 편리성을 향상시킬 수 있을 것으로 기대한다.

참고 자료

[1] 강호갑, "DRM 최신 국제표준 기술사양 분석 및 세계 유명제품 동향과 전망에 관한 연구", 2004,

한국소프트웨어진흥원

- [2] 강호갑/김태현, "DC 유통활성화를 위한 DRM 표준화 방법 연구", 2005, 한국소프트웨어진흥원
- [3] 강호갑, "콘텐츠 보호" (2005~2006 디지털콘텐츠산업백서, 2장), 2006, 한국소프트웨어진흥원
- [4] TTA DRM-PG WG1101, DRM 시스템과 이미지 평거프린팅 시스템의 통합 프레임워크, 2006, 한국정보통신표준협회
- [5] UCI 명세서 Version 2.1, 한국정보사회진흥원, 2006.