

논문 2009-3-12

P2P에서의 음악 콘텐츠 유통 시스템 설계

The Design of Circulation System of Digital Music Contents in P2P Service

강우준*

Woo-Jun Kang

요 약 본 논문은 P2P 서비스의 유료화를 효율적으로 수행하기 위해 필요한 음악 콘텐츠 유통 시스템을 제안한다. 제안 시스템은 콘텐츠 식별 방법, 저작권 DB 관리 방법, 요금 결제 방법 및 정산 방법 등을 제공한다. 일반적인 디지털 음악상품 유통 시스템에서는 운영자가 판매되는 음악 콘텐츠를 제공하지만, P2P의 경우는 사용자들이 음악 콘텐츠를 제공한다. 따라서 운영자가 유통되는 콘텐츠가 무엇인지 명확히 식별하는 것이 매우 어렵다. 일반적으로 콘텐츠 식별을 위해서는 해시를 사용하는 경우와 콘텐츠 식별 기술을 사용하는데, 본 논문에서는 해시 기술 및 콘텐츠 식별 기술을 융합하여 두 기술의 장점만을 활용함으로써 해시의 처리 속도와 콘텐츠 식별 기술의 편리함을 동시에 제공한다.

Abstract In this paper, we propose a new design of circulation system of digital music contents in P2P service. It has functions of contents recognition, management of copyright DB, clearing house to perform payment system of P2P service. In P2P service each user provides his music contents, but in other circulation system, service providers provide their music contents. Therefore it is very difficult for each provider to recognize whatever the quality of distributing contents are good or not. Most circulation systems use hash technology or contents recognition technology for the recognition of contents. For the usability and efficiency this proposed system provides the combining technology of hash processing and contents recognition.

Key Words : P2P, hash, contents recognition

1. 서론

인터넷 산업의 유료화가 확산되면서 안정적인 수익모델에 대해 많은 연구가 이어져 왔다[1]. 그럼에도 불구하고 디지털 음악상품의 합법적 이용과 관련하여 여전히 국내외적으로 많은 논란이 일고 있다. 특히 국내적으로 음반기획사, 신탁관리 단체, 메이저 음반제작사, 인터넷 음악서비스 사업자 등 시장 구성원들이 최근 P2P(Peer to Peer) 서비스 업체에 대한 '무제한 정액제' 허용을 놓고 이해관계의 차이로 서로 입장을 달리 하고 있기 때문이다. 현재 '무제한 정액제' 허용과 관련 디지털 음악 업

계에서는 "과도기적인 어쩔 수 없는 선택"이라는 음원 신탁관리 단체들의 주장과 "장기적으로 시장을 축소시키고 디지털 음악 산업 전체를 고사시킬 수 있는 근시안적인 사고"라는 대형음반사와 정규 음악서비스 업체들의 주장이 팽팽히 맞서면서 적잖은 갈등을 빚고 있다.

디지털 음악이 유통된 이후 음악서비스는 합법적인 저작권 관리 기술을 적용한 사이트 보다는 불법적인 P2P 서비스가 주를 이루어 왔다. 이와 같이 음악서비스를 주도하는 P2P 기술은 예외 없이 합법적인 서비스를 추구하기 보다는 법적인 공격이 강화될수록 저작권 회피의 수단을 강화하여 왔다. 그러나 음악 권리자들은 P2P 서비스 운영자뿐만이 아닌, 개인적인 음악 불법 제공자들에게도 법적인 처벌을 강화하고 있어 저작권 보호 및 합법

*정회원, 그리스도대학교 경영정보학부
접수일자 2009.5.23, 수정일자 2009.6.12

화는 피할 수 없는 추세라고 할 수 있다[2]. 또한 최근 IETF, ITU-T 등 표준화기구들에서 P2P 관련 표준화 활동이 점차적으로 증가되고 있는 추세이다[3]. 이러한 상황 속에서 유료화 모델은 P2P 서비스의 가장 중요한 이슈가 되고 있다.

본 논문에서는 P2P 서비스의 유료화를 효율적으로 수행하기 위해 필요한 콘텐츠 식별 방법, 저작권 DB 관리 방법, 요금 결제 방법 및 정산 방법 등을 제공하는 유료화 시스템의 설계를 제안하고자 한다.

과거에도 P2P 서비스들은 콘텐츠를 다운로드 할 때 주는 자와 받는 자간에 포인트를 주고받는 것이 일반화 되어 있었다. 그러나 현재의 이러한 방식은 다운로드 건별로 사용료를 지급하여야 할 권리자와 콘텐츠를 명확하게 연결할 수 있는 콘텐츠 식별 방법이 없는 상태에서 진행되는 것이다. 따라서 현재 진행되는 P2P 서비스에 있어 음악 분야 유료화의 최대 쟁점은 명확한 음악 식별 방법이라고 할 수 있다. 불특정 다수가 제공하는 콘텐츠는 일반적으로 정보가 부정확하고 일관성이 결여되어 있어 해시(Hash)나 음악 식별(Acoustic Fingerprint) 기술 등 신뢰할 수 있는 콘텐츠 식별 방법을 이용하여 저작권 관리를 하여야 한다[4].

이를 위해 제안 시스템에서는 해시 기술 및 콘텐츠 식별 기술을 융합하여 두 기술의 장점만을 활용함으로써 해시의 처리 속도와 콘텐츠 식별 기술의 편리함을 동시에 제공하고자 한다. 본 논문의 나머지 구성은 2장에서는 P2P 서비스의 동향을 알아보고 3장에서는 제안기술과 기존기술을 비교해 보았다. 4장에서는 제안 시스템의 구성에 대해 설명하였으며, 마지막으로 5장에서는 결론을 맺는다.

2. P2P 서비스의 기술 동향

2.1 P2P 서비스의 운영 형태

디지털 기술의 발전으로 말미암아 새로운 유형의 음악 서비스들이 속속 등장하고 있으며, P2P 음악서비스도 그 중 하나라고 할 수 있다[10]. 이러한 추세에 발맞추어 현재 운영되고 있는 다양한 음악서비스들은 물론이고, 향후 새로운 형태로 변화할 서비스들에 대하여 법적용이 용이하도록 규정을 개정할 바 있다. 개정 내용을 살펴보면, 기존에 있던 'P2P 음악서비스' 조항을 삭제하고, 음악

서비스 유형을 주문형 스트리밍, 주문형 다운로드, 주문형 배경음악으로 각각 분류하고 있다. P2P 음악서비스는 '주문형 다운로드 서비스'에 관한 조항을 적용받게 되며, 새로운 상품모델인 월정액 서비스에 대한 사용료 징수 규정도 신설하였다.

'음악저작물사용료징수규정' 개정에 따른 P2P 음악서비스의 변화 내용을 살펴보면, P2P 사이트 내 정산 데이터를 확보할 수 있도록 시스템을 구축하고, P2P 음악 서비스 운영을 통해 얻은 수익금 중 저작권자 등에게 지급하여야 할 비율에 따라 이익금을 분배하여야 한다고 규정하고 있다. 이와 같은 P2P 사이트의 유료화로 인해 무제한 다운로드, DRM 프리 등 새로운 비즈니스 모델들이 급성장하면서 불법 음원이 확산될 경로가 다양해질 가능성이 존재하는 것도 사실이다.

또한, 현행 저작권법 제 104조에서 규정하고 있는 특수한 유형의 온라인서비스제공자의 의무는 권리자의 요청이 있는 경우 당해 저작물 등의 불법적인 전송을 차단하는 기술적인 조치 등을 하도록 하고 있어 P2P 음악서비스에 사용 허락되지 않은 불법 음원이 유입되는 것에 대한 사전 방지 의무가 절실하다고 볼 수 있다. 새로운 음악 상품 유통 시스템을 융통성 있게 수용하면서 저작자의 권리 보호라는 원칙을 지킬 수 있도록, 양자 간의 적절한 조화를 이룰 수 있는 정책적인 방안 마련이 필요하다[5].

2.2 합법적 P2P 서비스의 기술적 요구사항

P2P 서비스가 합법적인 음악 서비스가 되기 위해서는 모든 음악은 누가 해당음악의 합법적 권리자인지를 확인할 수 있는 구조를 가져야 하며 저작권을 가진 음악은 반드시 보호된 음악의 형태로 배포되어야 한다. 모든 음악 관련 정보는 변조를 막을 수 있는 정품 인증 구조를 가져야 한다. 또한 충분한 양의 음원이 확보되어야 한다.

디지털 음악 상품의 권리 확인 구조는 첫째, 음악 권리자들이 그 안전성을 인정할 수 있는 수준이 되어야 하고 둘째, P2P 서비스 업체들에게는 기존의 장점 및 골격이 최대한 유지되어야 하며 셋째, 기존의 사용자들이 아무런 불편을 느끼지 않고 지속적인 참여가 가능한 수준이 되어야 한다. 넷째, 기존의 장점을 유지하기 위해 점진적으로 유료화를 하지만, 궁극적으로는 완전 유료화를 하여야 하며 합법적으로 음원 확보가 용이한 구조를 가져야 한다.

또한 음악서비스는 다음과 같은 안전성 관련 사항들이 고려되어야 한다. 첫째, 합법적인 이용비용이 불법적인 이용비용보다 낮게 만들어야 한다. 음악 상품의 가격이 낮아지면 상대적 안전성은 높아진다. 음악 상품의 가격이 높아진다고 하여도 그것이 매출의 증대를 가져오지는 않는다. 둘째, 사용성이 좋아야 한다. 사용성이 좋아지면 안전성이 증가한다. 또한 많은 요소들이 상대적 안전성에 영향을 미친다. 과도한 통제나 급격한 통제 강화는 음악서비스의 장애가 될 수 있다. 합법적인 음악서비스가 활성화되지 않은 상황에서는 낮은 가격 정책을 적용하는 것은 합법적인 음악서비스의 정착을 위해 불가피한 과정이라 할 수 있다[6].

더욱이 음원 ‘사용 허락’ 절차 없이 과금 체계만 적용함으로써는 기존 P2P 사업자들의 수익 구조개선 효과는 미비하고, 음원사업자들에 대한 저작권 보호 기능의 실효성도 기대할 수 없다. 음원 사용료 징수만으로 P2P 방식의 파일교환이 합법화 되었다고 할 수는 없다. P2P 사이트 내에서 교환되고 있는 음원들에 대한 불법성은 여전히 존재하고 있으며, 징수규정만 적용에 따라 P2P 사이트에서 이용자들에게 부과하는 음원 사용료가 인상될 시에는 징수규정을 적용하지 않는 불법적인 형태의 P2P 서비스 모델이 언제든지 생겨날 가능성도 있기 때문이다.

3. 기존기술과의 비교분석

3.1 콘텐츠 식별에 있어 기존기술의 문제점 분석

음악식별 기술은 조그만 음질 차이나 변형에도 동일 음악 여부를 잘 판별하여 저작권 관리를 쉽게 할 수 있도록 해주지만, 여전히 다음과 같은 문제점들이 있다.

첫째, 음악식별 기술은 해시 기술보다 훨씬 더 많은 자원이 필요하며 음악을 분석하는데 시간이 필요하므로 저작권을 조회하는 시간이 길어져 사용자의 불편을 증가시키고 서버 운용비용도 대폭 증가시키는 문제점이 있다. 따라서 이런 문제를 해결하기 위해 그동안 제시된 많은 기술들이 음악파일의 일부분만을 이용하여 식별하는 방식을 사용함으로써 식별의 정확률이 낮아지는 부작용을 초래하였다.

둘째, 콘텐츠 식별 정확률 문제가 있다. 사람이 듣고

동일한 음악으로 인정하는 것을 음악식별 기술이 모두 동일한 음악이라고 판단하지는 않는다. 또한 사람마다 음악을 듣고 동일한 음악이라고 판단하는 기준도 다르다. 실제로 리메이크나 라이브 음악의 경우 거의 같은 음악이지만 사실상 음악 인접권자가 다르고 사용료를 지급하여야 하는 대상도 다른 경우가 많다. 따라서 음악식별 기술만 가지고 100%의 식별 정확률을 얻는 것은 불가능하며 어느 정도의 사람의 개입이 필요하게 된다.

셋째, 음악 등록 방식에 대한 문제이다. 기존 음악 식별 기술은 서버 운영자가 음악에 대한 정보를 입력하고 이후 음악을 조회하는 측에서 부분 음악 특징을 가지고 이미 등록된 음악들과 일치되는 부분이 있는가를 검색하여 음악 정보를 얻는다. 그러나 이런 방식에서는 음악식별 실패 시 사후 대처를 하기 위해서는 음악식별이 실패된 음악을 다시 새로운 식별 대상으로 등록해야 한다. 새로운 식별 대상을 등록하는 가장 효율적인 방법은 자동으로 등록하는 것이며 수동으로 등록하는 것은 지나치게 많은 관리 비용이 들게 된다. 그러나 기존 음악식별 기술은 클라이언트에서 실제로 식별이 이루어질 때에는 음악의 부분적인 특징만을 사용하게 됨으로써 서버에 전체 음악에 관련된 정보가 전달되지 않으므로 클라이언트에 의한 자동 등록이 불가능하게 된다.

넷째, 기존 기술은 음악 가격이나 저작권 유형을 검색 단계에서는 알 수 없다는 문제가 있다. 이러한 문제점은 낮은 음악식별 기술의 성능이 그 원인이다. 저작권 조회를 하는 경우가 많으면 음악 식별의 처리 횟수가 증가하여 많은 비용이 들기 때문에 사실상 음악 다운로드 시점에만 음악식별 기술을 사용하여 구매 여부를 결정하고 있다. 이러한 방식은 많은 자원이 소요되는 음악식별을 꼭 필요한 구매 시점에만 적용하려하기 때문이지만 결정적으로 다운로드 이전에 가격이나 유료, 무료 여부를 알 수 없어서 사용성에 큰 문제를 발생시킨다. 가격과 유료 및 무료 여부는 음악 다운로드 이전의 검색단계에서 반드시 제공되어야 하는 정보이다[7].

3.2 콘텐츠 식별 DB 정보 입력 방법

기존의 콘텐츠 식별 기술은 서비스 운영자가 초기 콘텐츠 식별 DB 정보를 입력하는 것이 일반적이다. 그러나 본 제안 시스템에서는 불특정 사용자가 조회한 콘텐츠가 콘텐츠 식별 DB 항목에 들어 있지 않은 경우 조회 정보를 이용하여 해당 조회 정보 자체가 새로운 DB 항목이

됨으로써 효율적인 콘텐츠 식별 DB구축 방법을 제공한다. 콘텐츠 식별 실패 시 즉 같은 콘텐츠가 동일한 콘텐츠로 식별되지 않고 다른 콘텐츠로 식별되는 경우에 새로운 콘텐츠 그룹이 자동 등록되어 2개 이상의 콘텐츠 식별 DB항목을 연계시키면 식별 정확률을 크게 향상시킬 수 있다.

3.3 저작권 조회 시점

기존의 P2P 합법화 솔루션은 저작권 조회를 콘텐츠 교환 부분에 적용한다. 이는 공유 폴더의 모든 콘텐츠에 콘텐츠 식별 기술을 이용하여 저작권 조회를 하는 경우 많은 비용이 소요되기 때문이다. 그러나 본 제안 시스템에서는 해시 기술과 융합된 콘텐츠 식별 기술이 저작권 조회 당 비용을 최저 수준으로 하락시켜 저작권 조회를 콘텐츠 교환 부분뿐 만이 아닌 공유 콘텐츠 전체에 대해서 수행할 수 있게 하였다. 따라서 콘텐츠 교환 시에 저작권 유형을 결정하는 대신 콘텐츠 검색 시에 저작권 유형과 가격 정보를 제공하여 사용의 편의성을 획기적으로 개선하는 구조를 갖추고 있으며 음악 가격의 기본 값이 음질에 따라 자동 책정되는 기능도 제공한다.

3.4 음악의 전체 특징을 이용한 콘텐츠 식별

[그림 1]은 음악의 부분 특징을 이용하여 콘텐츠 식별을 하는 경우와 전체 특징을 이용하여 콘텐츠 식별을 하는 경우의 차이를 보여준다.

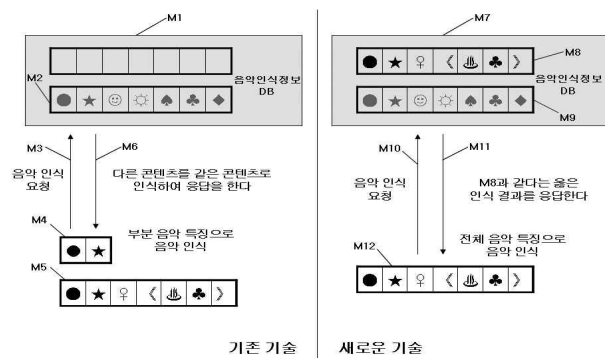


그림 1. 음악 식별을 위한 기존 기술과 제안 기술의 비교
Fig. 1 The comparison of existing and new technology in contents recognition

부분 특징을 이용하여 콘텐츠 식별을 하는 기존기술은 특정 콘텐츠(M5)의 부분 특징(M4)을 이용하여 서버에 음악식별 요청을 한다(M3). 요청의 결과는 실제로

M2와 M5가 다른 콘텐츠이므로 다르다고 해야 한다. 그러나 기존 기술은 부분 정보만 가지고 판단하므로 결국 특정 콘텐츠(M5)가 음악식별 정보 DB에 있는 콘텐츠(M2)와 동일하다는 잘못된 결과를 반환하게 된다(M6).

본 제안 시스템에서는 해시 융합으로 콘텐츠 식별 성능을 크게 개선시켰으므로 콘텐츠 식별을 직접 사용하는 상황에서는 전체 파일을 이용하여 이런 문제를 해결하였다. 즉 전체 파일(M12)을 이용하여 음악식별 요청(M10)을 하며 일부만 동일한 M9가 아닌 전체가 동일한 M8을 식별된 콘텐츠로 반환한다(M11).

또한 기존 DB상에 조회하는 음악과 동일한 정보가 없을 경우에 조회하는 정보 자체를 이용하여 DB를 확장할 수 있다. 새로운 기술의 M8은 기존에 있던 DB항목이 아닌 조회 당시에 등록된 내용이며 이런 방식으로 쉽게 DB항목이 확장되어 질 수 있다.

3.5 기존기술과의 비교분석 요약

기존의 디지털 음악상품 유통 시스템에서는 운영자가 판매되는 음악 콘텐츠를 제공하지만, P2P의 경우는 사용자들이 음악 콘텐츠를 제공하기 때문에 운영자가 유통되는 콘텐츠가 무엇인지 명확히 식별하는 것이 매우 어렵다.

표 1. 콘텐츠 식별의 효율성 비교
Table 1. The comparison of efficiency in contents recognition

	기존 기술	제안 기술
저작권 조회에 걸리는 시간(식별 기술의 경우)	느림	빠름
콘텐츠 식별 정확률	낮음	높음
콘텐츠 식별의 방법	콘텐츠의 부분만을 비교(식별 기술만을 사용)	콘텐츠의 전체를 비교(식별 기술과 해시 기술의 융합)
새로운 음악 등록 방식	서버 운영자가 등록	자동으로 등록
저작권 조회 시점	구매(콘텐츠 교환) 시점에 가능	구매 시점 이전에 언제든지 가능
콘텐츠 식별 DB 정보 입력(조회한 콘텐츠가 DB에 없는 경우)	운영자가 서비스 초기에 실시	새로운 콘텐츠 발견 즉시 자동으로 등록

콘텐츠 식별을 위해서는 일반적으로 해시를 사용하거나 콘텐츠 식별 기술을 사용하는데, 본 제안에서는 해시 기술 및 콘텐츠 식별 기술을 융합하여 두 기술의 장점만을 활용함으로써 해시의 처리 속도와 콘텐츠 식별 기술의 편리함을 동시에 제공한다.

[표 1]은 콘텐츠 식별에 있어서 기존 기술과 제안 기술을 비교분석한 결과를 요약한 것이다.

해시는 처리 속도 측면에서 유용하지만 무의미한 내용의 변화에도 값이 달라져 많은 수의 해시 값을 관리하여야 한다. 콘텐츠 식별 기술은 관리 대상인 콘텐츠의 개수를 최소화하여 운영의 효율성을 제공한다.

3.6 제안 기술의 성능실험 결과

본 제안 시스템은 상용화를 목적으로 설계되고 구현을 하였으므로 실제적인 테스트 환경에서 실제의 음악들을 대상으로 실험을 실시하였다. 그 결과 [표 2]와 같은 실험 결과를 얻을 수 있었다.

표 2. 음악 식별 정확률 및 소요시간 측정 결과
Table 2. The test results of accuracy and time required in contents recognition

	실험 결과 값
정확률	95 % (1754 - 97) / 1754 = 0.95 실험결과 97 개 파일 그룹핑 에러
서버에서 음악식별 소요시간	6 milli-second (10.544 sec / 1754 = 0.006 sec)
생성 그룹의 수	193 개
* 실험 대상 - 음악: 55 개의 사용 음악(곡) - 생성 해시 값: 1,754 개 파일 - 그룹의 수: 같은 곡이지만 약간의 변형 등이 원인이 되어 다른 그룹으로 분류된 경우도 포함	

4. 제안 시스템

4.1 제안 시스템의 구성도

P2P의 가장 기본적인 형태에 있어서, 배포되는 콘텐츠는 네트워크 상에서의 파일의 제작, 검색 등이 효율적으로 이루어져야 한다[8][9].

제안 시스템에서는 저작권 조회 단계, 저작권 조회 서버에 새로운 내용을 적재하는 단계, 조회 결과를 반환하

는 단계, 조회 실패 시 저작권 항목 등록 준비 단계, 콘텐츠 패턴 ID 값을 얻어 오는 단계, 저작권 항목 등록을 하고 저작권 조회를 완료하는 단계를 포함하는 저작권 정보 조회/등록 시스템이 제공된다.

또한 판매 정보 저장 단계, 판매 보고서 생성 단계, 사용자 정산 정보 생성 단계를 포함하는 콘텐츠 정산 시스템이 제공되며, 변경 대상 콘텐츠 정보 검색 단계, 저작권자 정보 입력 단계, 해시 기반 정보 생성 단계를 포함하는 저작권 정보 입력 시스템이 제공된다.

여기서 저작권 정보 조회/등록 시스템과 저작권 정보 입력 시스템은 저작권 정보 관리 시스템의 일부이며 저작권 조회 시에 자동으로 처리되는 부분과 사람이 직접 제공할 수밖에 없는 정보를 입력하는 부분으로 구분된 것이다.

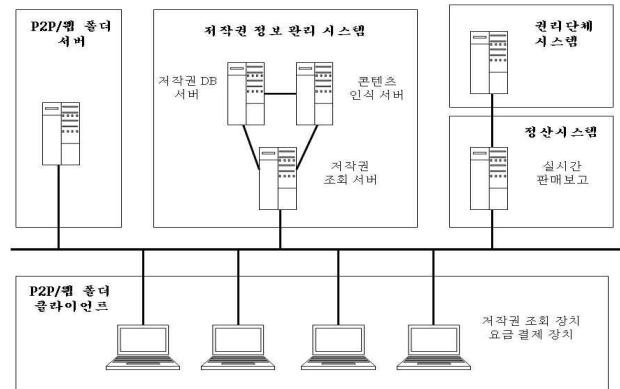


그림 2. 시스템 구성도
Fig. 2 The configuration of proposed system

[그림 2]는 본 논문에서 제안하는 시스템 구성도이다. 기존 P2P 서비스 기능에 콘텐츠 식별 수단을 제공하고, 이 정보를 바탕으로 저작권 정보를 얻어 이에 맞게 처리하고 판매 결과를 권리단체에 정산하는 구조를 보여 준다. P2P 서버와 클라이언트는 기존 P2P 서비스가 가지고 있는 기능이며, 제안 시스템에 포함되는 API 프로그램을 삽입하여 유료화가 가능한 형태로 수정될 수 있다. 그 밖의 저작권 DB 서버, 저작권 조회 서버, 콘텐츠 식별 서버, 정산 시스템, 권리 단체 서버 등은 제안 시스템이 제공하는 부분이다.

기존 서비스의 기능에 유료화를 위해 추가되어야 하는 것은 저작권 조회 및 판매 보고 기능이다. 이 두 기능이 P2P 서비스의 경우에는 클라이언트에 추가되는 것이 일반적이다.

4.2 저작권 조회 서버

제안 시스템에서는 저작권 조회의 위치와 관련 없이 모든 공유 콘텐츠에 대해 저작권 조회가 가능하게 한다. 기존기술은 콘텐츠 식별 기술의 성능이 매우 낮아 저작권 조회 발생 수를 줄이기 위해 구매가 발생하는 콘텐츠 교환 단계에만 저작권 조회를 허용하지만, 본 시스템에서는 모든 공유 콘텐츠에 대해 저작권 조회를 수행할 수 있으므로 검색 단계에서부터 저작권 정보를 이용 가능하게 한다.

저작권 정보에 대한 조회가 발생하면 저작권 조회 서버가 이를 처리한다. 저작권 조회 서버는 해시 기반으로 저작권 정보가 메모리 캐싱이 된 서버로서 한 개 서버의 초당 처리 건수가 만 건이 넘으며 캐싱이 주 역할이기 때문에 손쉽게 여러 대로 확장이 가능하다. 실제로 높은 성능을 필요로 하는 경우에는 여러 대의 조회 서버를 설치할 수 있다. 기능상 저작권 조회 서버는 해시 기반의 서버이며, 저작권 DB 서버는 해시와 콘텐츠 식별 기술이 융합된 구조를 갖고, 콘텐츠 식별 서버는 콘텐츠 식별 기능을 기반으로 작동된다. 저작권 DB 서버는 해시 기반 정보와 콘텐츠 패턴 ID 기반 정보를 모두 소유하는 서버이다. 이와 같이 해시 정보와 콘텐츠 패턴 ID 정보가 융합되어 있으므로 해시 단위의 빠른 처리가 가능하고, 콘텐츠 패턴 ID 단위의 편리한 저작권 정보 관리가 가능하다.

저작권 조회 서버에 새로운 내용을 적재하는 단계는 메모리 캐싱을 이용하여 성능을 향상 시키는 역할을 하는 저작권 조회 서버가 최근의 저작권 정보를 가지고 있지 않을 때 최근의 정보를 적재하는 단계이다.

4.2.1 저작권 조회 위치

[그림 3]은 저작권 조회를 하는 위치를 보여 주기 위한 도면이다. C1은 주는 측의 공유 폴더이며 대부분의 저작권 조회가 발생하여 미리 저작권 정보를 보유한 상태가 된다.

따라서 검색 요청(C2)에 저작권 유형이나 가격 정보를 검색 결과(C3)에 넣어 줄 수 있다. 기존 기술은 검색 결과에 저작권 유형이나 가격 정보가 포함될 수 없으며, 이는 사용자들에게 불편을 주게 된다. 검색 후에 파일이 다운로드 되면(C4) 임시 저장소에 저장되며, 기존 기술들은 이곳에서만 저작권 조회를 수행한다. 이곳에서 요금 지급이 완료되면 정식 저장소인 고유폴더(C6)로 이동된다. 기존 기술은 콘텐츠 식별 기술의 사용 횟수를 줄이기

위해 C5에서 저작권 조회를 하지만, 제안 시스템에서는 모든 위치에서 모든 공유 콘텐츠에 대한 저작권 조회를 할 수 있다.

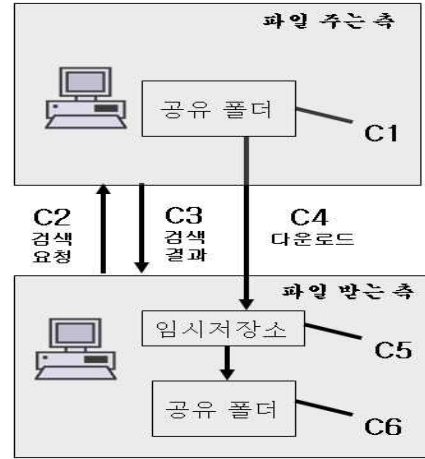


그림 3. 저작권 조회 위치
Fig. 3 The location of the inquiry of digital rights

4.2.2 저작권 정보 DB의 구성도

[그림 4]는 저작권 정보 DB의 개념적 구성도이다. 해시 정보와 콘텐츠 식별 기술로 생성된 정보가 융합하여 관리되는 구조를 보여준다.

해시 값	해시 단위 정보	콘텐츠 패턴 ID 값	콘텐츠 패턴 ID 값 단위 정보
해시 1	가격정보, 음질, ...	콘텐츠 패턴 ID 1	저작권자 정보 1, 저작권자 정신품 게재 정보 1, 등
해시 2	가격정보, 음질, ...		
해시 3	가격정보, 음질, ...		
해시 4	가격정보, 음질, ...	콘텐츠 패턴 ID 2	저작권자 정보 2, 저작권자 정신품 게재 정보 2, 등
해시 5	가격정보, 음질, ...		
해시 6	가격정보, 음질, ...		
해시 7	가격정보, 음질, ...	콘텐츠 패턴 ID 3	저작권자 정보 3, 저작권자 정신품 게재 정보 3, 등
해시 8	가격정보, 음질, ...		

그림 4. 저작권 정보 DB의 개념적 구성도
Fig. 4 The conceptual table of DB of digital rights information

해시 관련 정보는 콘텐츠 패턴 ID 값 관련 정보에 비해 레코드 건수가 훨씬 많다. 따라서 콘텐츠 패턴 ID 값을 기준으로 저작권 조회/등록을 하면 전체 건수 측면에서 여러 가지 장점을 갖지만 처리 속도 측면에서는 매우 느리다는 단점을 갖는다. 그러므로 두 가지 정보가 그림 4와 같이 연계되면 상황에 맞게 이용할 수가 있다. 저작권 DB 서버는 해시 기반 정보와 콘텐츠 패턴 ID 기반 정

보 모두를 보유하는 서버이다. 이와 같이 두 정보가 융합되어 있으므로 해시 단위의 빠른 처리가 가능하고 콘텐츠 패턴 ID 단위의 편리한 저작권 정보 관리가 가능하다.

4.2.3 저작권 조회 과정 순서도

[그림 5]는 해시와 콘텐츠 식별 기술을 융합하여 저작권 조회가 효율적으로 처리되는 과정을 보여주는 순서도이다. 기존의 기술들은 운영자의 자료 제공 없이는 처리가 시작되지 못하고 콘텐츠 식별 실패 등의 문제가 발생하는 경우에도 운영자의 개입이 없으면 조회 실패로 처리가 끝나는 불완전한 구조로 되어 있으나, 제안 시스템에서는 사람의 개입 없이 모든 것이 자동으로 수행될 수 있다.

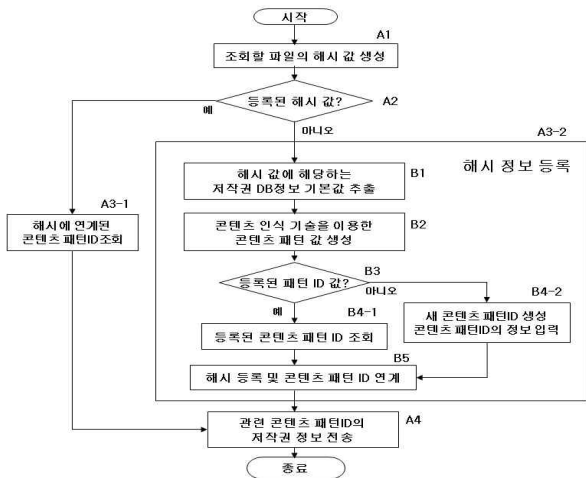


그림 5. 저작권 조회 과정 순서도
Fig. 5 The flow chart of inquiry process of digital rights

또한, 기존 기술들은 속도가 빠른 해시 방법과 콘텐츠 식별 기술이 융합되어 있지 않아 처리 속도가 매우 느리고 운영비용이 많이 소요되나, 제안 시스템에서는 자동으로 해시와 콘텐츠 식별 기술이 융합되는 과정을 보여준다. 저작권 조회(A1)는 해시 값을 기반으로 수행된다. 해시 값 처리 속도는 콘텐츠 식별 기술보다 클라이언트 측에서 20-70배 정도 빠르며, 서버 측에서는 100-1000배 빠르다. 따라서 콘텐츠 식별 기능의 사용이 얼마나 많은가 여부가 전체 운영비용을 결정짓는다고 할 수 있다. 제안 시스템에서는 모든 해시 값에 콘텐츠 패턴 ID를 연계시키기 위한 목적으로만 콘텐츠 식별 기능을 사용하기 때문에 콘텐츠 식별 기능의 사용을 최소화하였다.

현재 전체 음악 해시 값의 개수는 1200만개로 추산하

고 있으며, 매일 새로 출현하는 해시 값은 10만 개 이하로 추산된다. 전체 음악 파일 개수를 10억 개로 계산하고 저작권 정보 변경 때문에 매일 새로 저작권 조회를 수행하는 것을 가정하면, 하루당 해시 기반으로 처리되는 저작권 조회는 10억 건, 음악 식별 기능이 호출되어 새로운 해시 등록 절차가 수행되는 것은 매일 10만 건으로 계산할 수 있다. 기존 기술에서 저작권 조회를 콘텐츠 식별 기능을 이용하여 처리하려고 한다면 이는 10배 내지 1000배의 비용을 감수해야 한다. 이러한 이유 때문에 기존 기술은 저작권 조회 자체를 최소화하기 위해 구매가 발생하는 콘텐츠 다운로드 경우에만 저작권 조회를 허용하는 방식을 적용하게 되며, 따라서 교환이 진행되지 않으며, 유료/무료 여부, 가격 등을 알지 못하는 커다란 불편을 감수해야 한다.

A2에서 저작권 조회 서버는 기 등록된 해시 값인 지 여부를 조사하여 기 등록된 값이면 해당 테이블에서 저작권 정보를 얻어 클라이언트에게 반환한다(A3-1, A4). 등록되지 않은 값이면 해시 정보 등록을 수행한다(A3-2). 클라이언트는 해시 값 등록에 필요한 정보들을 콘텐츠로부터 추출하고(B1) 콘텐츠 식별 기능 수행에 필요한 콘텐츠 패턴 값도 생성하여(B2) 저작권 DB 서버에 등록 요청을 한다. 저작권 DB 서버는 콘텐츠 식별 서버에 해당 콘텐츠 패턴 ID 값을 요청한다. 콘텐츠 식별 서버는 해당 콘텐츠 패턴 값이 소속할 수 있는 콘텐츠 패턴 ID 값이 있으면(B4-1) 이를 반환하고, 없으면 새로 콘텐츠 패턴 ID 값을 생성하여(B4-2) 이를 반환한다. 콘텐츠 패턴 ID 값을 얻으면 해시 값, 해시 관련 정보, 콘텐츠 패턴 값, 콘텐츠 패턴 ID 값, 콘텐츠 패턴 ID 관련 정보가 연결되어 해시 값, 콘텐츠 패턴 값 또는 콘텐츠 패턴 ID 값 어느 것을 이용하여도 관련 정보를 얻을 수 있는 구조가 완성된다(B5). 이 값들은 저작권 DB 서버에 저장된다. 이러한 해시 등록 과정이 끝나면 해당 해시에 연결된 콘텐츠 패턴 ID 값 관련 정보인 저작권 정보를 클라이언트에 반환한다(A4).

4.3 정산 시스템

정산 시스템은 P2P 서비스에서 판매가 발생하면 판매 보고를 받아 처리하는 기능을 수행하는 서버이다. 이 서버는 모든 판매 자료를 저장하고 이를 저작권자 별로 정리하여 권리자 또는 권리자 단체에 판매 보고 및 정산을 수행한다.

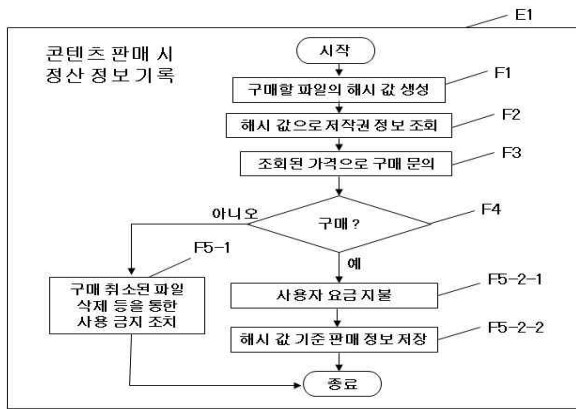


그림 6. 콘텐츠 판매 시 정산 정보 기록 순서도
 Fig. 6 The flow chart of the recording of payment information in the circulation of contents

권리 단체 시스템은 권리 단체의 서버들로 정산 서버로부터 판매 자료를 넘겨받는다.

저작권 정보 조회/등록 시스템의 첫 단계인 저작권 조회 단계는 콘텐츠 파일의 해시 값을 생성하고 이 값을 이용하여 진행된다. 모든 해시 값은 처음에 반드시 콘텐츠 식별 기술에 의해 만들어진 콘텐츠 패턴 ID와 연계되어 DB에 저장되므로 일단 한 번 사용된 해시 값을 갖는 콘텐츠는 다시 콘텐츠 식별 기술을 사용하지 않도록 방지되어 있다.

[그림 6]은 콘텐츠 정산 시스템의 일부로서 콘텐츠 판매 시 정산 정보를 기록하는 순서도이다. 이러한 처리 절차는 속도가 빠른 해시 값을 기준으로 처리되어 효율을 극대화시킬 수 있다. 해시 값은 저작권자별로 자료를 정리하기 위해 반드시 필요한 콘텐츠 패턴 ID 값이 연계되어 있어 전체적인 처리 절차는 아무 불편을 가져오지 않고 자동으로 처리가 된다. 구매가 발생하면 구매한 파일의 해시 값을 얻어(F1) 저작권 정보를 조회하고(F2) 조회된 콘텐츠의 가격을 이용하여 사용자에게 구매 여부를 물어본다(F3, F4). 구매를 거부하면 다운로드 받은 콘텐츠를 삭제하여 사용금지 조치를 취한다(F5-1). 구매를 하면 요금 결제를 한 후(F5-2-1) 해시 값 기준의 판매 정보를 저장한다(F5-2-2).

다음 단계는 [그림 7]에서 계속 된다. 그림 7은 정산 시스템의 순서도이며, 해시 기준으로 저장된 판매 정보가 저작권자별 판매 자료로 정리되는 과정을 설명한다. 그림6은 해시 기반으로 판매 정보를 저장하는 절차를 보여주었으며 이것은 그림7의 E1에 해당한다. 해시 기준으

로 저장된 판매 정보는 해시에 연계된 콘텐츠 패턴 ID 값을 이용하여 저작권자 정보를 얻고(E2) 콘텐츠 패턴 ID 값 기준으로 집계된(E3) 금액은 곧 저작권자별로 집계가 된다(E4). 이 저작권자별 정보는 추후 정산용으로 사용된다(E5).

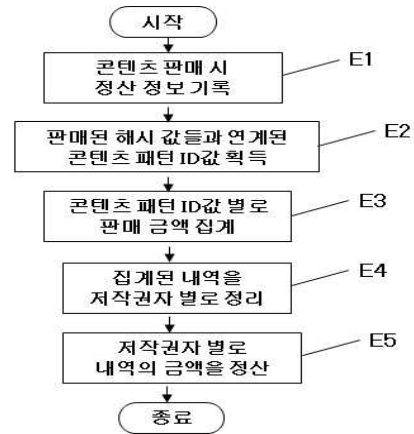


그림 7. 콘텐츠 정산 시스템의 순서도
 Fig. 7 The flow chart of payment system of the circulation of contents

5. 결론

본 논문에서는 P2P 서비스의 유료화를 효율적으로 수행하기 위해 필요한 콘텐츠 식별 방법, 저작권 DB 관리 방법, 요금 결제 방법 및 정산 방법을 제안하였다. 이를 위해 처리 속도가 탁월한 해시 기능과 저작권 정보 관리를 용이하게 해주는 콘텐츠 식별 기술을 융합하여 처리 속도의 문제없이 저작권 조회, 저작권 정보 입력 및 정산 등을 처리할 수 있는 시스템 구성을 설명하였다.

저작권 정보 관리 노력은 음악 식별 기술을 사용하면 약 200만 건에 대한 레코드 관리가 필요했지만, 해시 기술만을 사용하면 1200만 건 이상의 레코드 관리가 필요하다. 제안 시스템에서는 해시의 처리 효과를 가지면서도 200만 건의 레코드만 관리하면 되도록 설계하였고 또한 콘텐츠 식별 기술의 속도를 해시 수준으로 만드는 효과가 있으므로 저작권 조회에 제약을 주지 않고 아무 곳에서나 조회를 할 수 있게 하였다. 즉 기존기술에서 콘텐츠를 다운받아야 유·무료 여부를 알 수 있어서 사용의 편의성에 커다란 제약을 주었던 문제를 해결함으로써 사용의 편의성을 향상 시켰다.

콘텐츠 식별 기술 자체에도 질적 향상을 제고하였는

데, 서버에서 운영자가 콘텐츠를 등록하는 하향식 방식의 경우 운영자가 미리 등록을 하지 않았으면 콘텐츠 조회 실패가 발생하지만, 제안 시스템에서는 새로운 콘텐츠가 출현 시 자동 등록되는 상향식 방식에서는 콘텐츠 식별 실패가 발생할 수 없으며 새로운 콘텐츠가 출현하여도 기본 값이 자동 등록되어 저작권 조회 실패가 발생하지 않는다. 실제로 정산 시점에 필요한 개별 저작권자 정보 입력 등을 제외하고는 모든 저작권 관리가 자동으로 수행된다.

향후 연구에서는 큰 파일의 처리 속도를 향상시키는 방안을 마련하고 음악 파일이 아닌 동영상 파일 등의 콘텐츠 식별 기술에 대한 연구를 수행할 예정이다.

참고문헌

- [1] 박정서, “인터넷사업의 유료화 사례분석 및 확산방안 연구”, 정보통신 일반정책 지정공모과제, pp. 132-141, 2002.
- [2] 윤석찬, “인터넷저작물에 있어 저작권침해와 저작권법”, 세창출판사, pp. 117-132, 2002.
- [3] Jae Hoon Nah, Hyeok Chan Kwon and Jong Soo Jang, Proposal for studying P2P Network Security, D78, ITU-T, Oct., 2005.
- [4] 진명희, 이승은, 손영성, 김정식, “DHT 기반 P2P 시스템에서 키워드 검색 지원을 위한 시스템 디자인”, 한국정보과학회 가을 학술발표논문집, 제 31권, 제2호(III), pp. 550-552, 2004.
- [5] 황근수, “인터넷 상의 저작권보호 및 저작권법 제25조의 문제점”, 법학연구 제29집, pp. 499-522, 2008.
- [6] 강우준, 김응모, “디지털 저작권 관리 기술을 이용한 MP3 디지털 음악의 온라인 유통”, 한국정보처리학회 논문지 A, Vol. 07, No. 11, pp. 3694-3701, 2000.
- [7] 성재연, 정연정, 윤기승, “콘텐츠 보호를 위한 DRM이 적용된 P2P 모델”, Vol. 13-C, No. 04, pp. 0389-0396, 2006.
- [8] S. Androutsellis-Theotokis and D. Spinellis, “A Survey of Peer-to-Peer Content Distribution Technologies,” ACM Computing Surveys, Vol. 36, No. 4, 2004.
- [9] Reuning, J. and Jones, P., “Peer-to-Peer enabled content distribution”, Proceedings of the 5th ACM/IEEE-CS Joint Conference, pp. 396-396, 2005.
- [10] 이일우, 박호진, “P2P 서비스 응용 및 과금 기술 동향”, 전자통신동향분석, 제22권, 제5호, pp.121-129, 2007.

※ 이 논문은 2008년도 그리스도대학교 학술연구비 지원에 의한 논문임.

저자 소개

강 우 준(정회원)



- 1984년 2월 : 연세대학교 공과대학 전자공학과 (공학사)
- 1984년 1월 ~ 1999년 2월 : 한국 IBM 소프트웨어 연구소
- 1992년 8월 : 연세대학교 산업대학원 전자계산학 전공 (공학석사)
- 1995년 2월 : 연세대학교 경영대학원 MIS 전공 (경영학석사)

- 1999년 3월 ~ 2001년 2월 : 안산공과대학 컴퓨터공학과 교수
- 2001년 8월 : 성균관대학교 공과대학원 전기전자 및 컴퓨터공학 전공 (공학박사)
- 2001년 3월 ~ : 그리스도대학교 경영정보학부 교수

<관심분야> 콘텐츠 인식, DRM, 전자상거래 보안, 접근 제어, XML/Web 마이닝