

# 개념지도 기반 P2P 콘텐츠 공유 시스템

나고운<sup>o</sup> 차재혁

한양대학교 정보통신공학과

yepyp@ihanyang.ac.kr<sup>o</sup>, chajh@hanyang.ac.kr

## Concept Map based P2P contents sharing system

Na, Go-Un<sup>o</sup> Jae Hyuk Cha  
Hanyang University

### 요 약

현재의 P2P 시스템은 특정한 분야에서는 성공을 거두고 있지만 지식과 정보 공유의 목적으로서는 활용되지 못하고 있다. 사용자가 웹에서 수집한 자료와 새로 생성한 지식 및 정보 콘텐츠들은 기존의 P2P 시스템으로는 공유하기 힘든 몇 가지 이유가 있다. 그 이유 중 하나는 대부분의 P2P 시스템은 사용자가 콘텐츠 각각에 대한 메타데이터를 작성하도록 하거나 키워드가 나타도록 파일명을 수정하도록 요구하기 때문이다. 본 논문에서 제안하는 CONCEPT P2P는 P2P 환경에서 콘텐츠를 공유하는데 개념지도를 사용 한다. 개념지도는 시간 소모적인 메타데이터 작성 작업에 대한 대안이 될 수 있기 때문이다. CONCEPT P2P는 로컬 파일 시스템에 수집하고 있는 관심 콘텐츠의 공유를 목적으로 한다. 사용자의 개념지도를 추출하여 개념지도로부터 콘텐츠 묶음을 검색하므로 무리한 수작업이 없이도 다양한 분야의 콘텐츠를 효과적으로 공유할 수 있도록 설계 하였다. 또한 다른 사용자와 개념 네트워크를 형성하여 관심 개념에 의해서 클러스터링 되고 있는 복수개의 콘텐츠를 지속적으로 수집하고 배포하는 접근을 취하였다. 개념 지도를 기술하고 검색하는 부분에서는 RDF 기반의 시맨틱 웹 언어와 RDF 질의 언어인 QEL을 사용하였다.

### 1. 서 론

퍼스널 컴퓨터가 보급되면서 일반 사람들도 지식을 표현하고 보관하는 수단으로서 컴퓨터를 사용하는데 익숙해지게 되었다. 또한 인터넷의 확산과 대용량 하드디스크의 가격 인하로 인해 개인이 인터넷으로부터 콘텐츠를 수집하고 그 자료를 하드디스크에 오랜 기간 보관해 두는 것은 자연스러운 일이 되었다.

만약 각각의 퍼스널 컴퓨터가 거대한 콘텐츠 공유 네트워크를 형성하게 된다면, 보다 많은 지식과 정보를 공유할 수 있을 것이다. 본 논문에서는 로컬 파일 시스템에 수집한 콘텐츠의 효과적인 공유를 위한 시스템의 설계와 구현에 대한 내용을 다루고자 한다. 한정된 분야에서만 활용되고 있는 P2P 콘텐츠 공유 시스템을 지식과 정보 공유의 용도로서 확장하여 사용하기 위해서는 새로운 형태의 콘텐츠 공유 방식이 요구된다.

기존의 P2P 콘텐츠 공유 시스템에서 콘텐츠 공유는 콘텐츠 검색을 통해 이루어진다. 사용자는 필요할 때마다 키워드를 입력하여 콘텐츠를 찾고 다운로드 받는다. 그런데 콘텐츠가 제대로 검색 되도록 하기 위해서는 매우 번거로운 작업이 사용자에게 요구된다. 모든 콘텐츠에 메타데이터를 작성하거나 파일명에 키워드가 나타나게 파일명을 일일이 수정해야하는 것이다. 이러한 이유로 인해서 P2P 환경에서의 콘텐츠 공유 시스템은 지식이나 정보 공유의 용도로 쓰이기보다는 매우 한정된 분야에 머무르고 있다.

본 논문에서는 개념지도 기반의 P2P 콘텐츠 공유 시스템을 제안한다. 제안하는 P2P 시스템은 콘텐츠 메타데이터 검색 중심이 아니다. 사용자가 관심개념을 만들면 그것에 따라 연결된 또 다른 사용자의 개념들과 연결 관계를 만들어서 지속적인 콘텐츠의 수집과 배포를 통해서 콘텐츠를 공유한다. CONCEPT P2P는 사용자의 관심 개념에 의해서 클러스터링 된 복수개의 개념과 콘텐츠를 지속적으로 수집하고 지속적으로 배포하는 것에 초점을 두고 있다.

### 2. 배경 지식 및 관련 연구

#### 2.1 개념지도

개념 지도(Concept Map)는 개념과 그 개념들의 관계로서 표현하여 지식을 표현하고 조직화하기 위한 도구로서 사용된다. 개념지도를 계층구조로 표현할 때는 가장 일반적인 구조를 상위 레벨에 표현하고 그것보다 덜 일반적인 구조를 하위 레벨에 표현한다.[1]

CONCEPT P2P는 개인이 관심있는 콘텐츠를 수집하여 자신의 파일 시스템에 정리할 때 그 디렉토리 구조가 지식이 확장되어감에 따라 개념지도와 비슷한 형태로 확장되는 점에 착안하여 디렉토리 구조로부터 개념지도를 추출하는 방법을 취하고 있다.

#### 2.2 시맨틱 웹

W3C가 중심에서 활동하고 있는 시맨틱 웹[2]은 잘 정의된 의미를 통해서 데이터를 보다 더 잘 공유하고 재활용하기 위한 공통적인 프레임워크를 제공한다. 시맨틱웹은 RDF(Resource Description Framework)를 근간으로 하고 있다. RDF[3]는 웹 자원에 대한 정보를 설명하기 위한 언어이다.

본 시스템은 개념지도 데이터베이스를 만들고 개념지도를 활용하여 검색하기 위해서 RDF를 기반으로 개념지도를 기술한다.

#### 2.3 P2P

P2P(Peer-to-Peer)는 네트워크 상에 있는 모든 장치들간에 직접적으로 통신하는 모델이다. P2P는 특정한 아키텍처나 기술 이라기보다는 모든 유형의 파일 공유, 콘텐츠 배포와 전달에 대한 새로운 유형, 인스턴스 메시징, 분산 검색, CPU 또는 저장 자원의 공유 등과 같은 것들을 가능하게 하는 모델이다. [5]

SUN에서는 이러한 P2P 시스템간의 상호운용성을 보장하기 위해 JXTA[6]라는 P2P 프로토콜을 개발하고 있다.

EDUTELLA[7]는 SUN의 JXTA에 기반하고 있는 P2P시스템 중의 하나이다. EDUTELLA는 분산 메타데이터 저장소를 구축하기 위한 오픈 P2P 프레임워크이다. 이 시스템은

QEL(Query Expression Language) RDF 질의 언어를 사용해서 저장소의 정보를 검색하고 탐색하며 고수준의 메타데이터 처리를 시도한다.

CONCEPT P2P 에서는 개념지도와 콘텐츠를 공유하기 위해서 EDUTELLA P2P 네트워크에 연결 하며 EDUTELLA에서 제안한 QEL-RDF를 사용하여 질의를 작성하여 P2P 네트워크로 보내고 그 결과를 받아온다.

### 2.4 콘텐츠 신디케이션

콘텐츠 신디케이션이란 콘텐츠의 전체나 일부를 다른 사이트에서 볼 수 있도록 제공하는 서비스이다. [8] XML 신디케이션 포맷으로는 RSS 0.91, RSS 0.92, RSS 2.0, RSS 1.0 등이 있다. 이러한 포맷을 사용하여 새로 생성된 콘텐츠의 전체 정보 혹은 요약정보를 효과적으로 신디케이션 할 수 있다.

콘텐츠의 유통에서는 콘텐츠의 개념 구조뿐만 아니라 콘텐츠의 신선도도 중요하게 다루어져야 할 가치이다. 복잡한 디렉토리 구조를 가진 저장소에서는 새로 생성된 콘텐츠가 무엇인지 알기 어렵다. 파일시스템 디렉토리를 기반으로 하는 CONCEPT P2P 에서는 디렉토리 구조의 복잡성 때문에 새로 생성되는 콘텐츠가 묻혀버리는 것을 방지하고자 신디케이션 언어를 사용하여 웹에 퍼블리싱한다.

### 3. 시스템 사용 시나리오

CONCEPT P2P는 일반 사용자의 파일 시스템 디렉토리에 저장해 둔 콘텐츠를 사용한다. 관심 분야에 따라 클러스터링 된 콘텐츠를 지속적으로 수집하고 배포할 수 있도록 하여 지식과 정보의 유통 환경으로서 P2P 네트워크를 사용하는 것이다.

CONCEPT P2P는 크게 콘텐츠 수집과 배포의 기능을 가지고 있다. CONCEPT P2P는 사용자의 개념 공유에 의한 콘텐츠 공유 방식을 가지고 있다.

#### 3.1 콘텐츠 수집

A는 최근 이슈가 되고 있는 'ADL SCORM' 이라는 생소한 개념에 관심이 생겨 그것과 관련된 연구 자료를 수집하고자 한다. 또한 국내에 비슷한 관심을 가진 사람들이 누가 있으며 그 사람들은 어떤 콘텐츠를 통해 지식과 정보를 얻고 있는지도 궁금하다.

A는 CONCEPT P2P 시스템을 구동하여 자신이 잘 알고있는 전문가의 이메일이나 ADL 대표 사이트, 혹은 SCORM과 관련된어있는 키워드를 넣어 유사한 관심을 가진 다른 사용자 리스트를 얻는다.

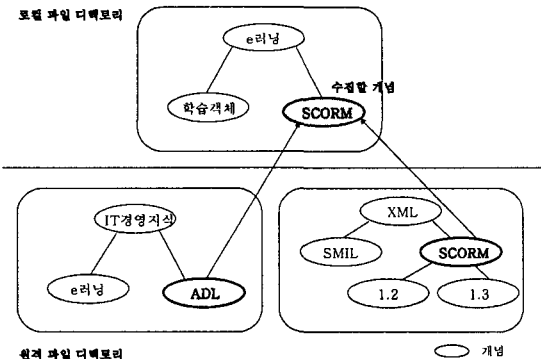


그림 1 다른 사용자와의 개념 공유

그리고 여러 사용자들의 개념지도도를 확인한다. 시각화된 인터페이스를 통해서 B라는 사람은 이러닝 비즈니스 모델과 연관된 자료를 모으고 있고 C라는 사람은 교수설계에 관심이 많으며 D라는 사람은 XML과 관련된 자료를 모으고 있다는 사실을 짐작할 수 있다. A는 자신의 SCORM이라는 디렉토리 안에 B의 ADL 디렉토리와 D의 SCORM 디렉토리를 연결한다. 그리고 A는 수시로 SCORM 디렉토리를 열어서 자동으로 다운로드 된 콘텐츠를 확인하여 학습하고 더 상세화된 콘텐츠 분류체계를 만들어 지식 영역을 확장한다.

### 3.2 콘텐츠 배포

A는 e러닝과 관련된 프로젝트를 수행하면서 매우 방대한 자료를 수집하게 되었다. A는 최신 자료를 자신의 파일 디렉토리에 업로드 할 때 마다 그 자료를 묶어서 관심 있는 사람에게 자료집을 배포하는 작업을 한다. A는 학습설계와 관련된 좋은 자료는 B에게 e메일로 알려주고 저장소와 관련된 자료가 나오면 C사의 FTP에 업로드해 준다. 그리고 개념이 쉽게 설명된 자료가 나오면 W사이트 자료실에 업로드 한다.

A는 매우 번거로운 작업이 반복됨에 따라 FTP 서비스를 구동하고 콘텐츠를 다운받을 수 있도록 하였다. 그런데 이것으로 인해서 '매우 복잡한 디렉토리 구조로 되어있어서 어디에 무슨 콘텐츠가 업데이트 되고 있는지, 내가 관심있는 콘텐츠가 어디에 있는지 모르겠다.'는 불만이 터져 나오게 되었다.

A는 CONCEPT P2P 시스템을 구동하여 자신의 디렉토리 몇 개를 선택하여 P2P 공유를 허락하였다. 그러자 시스템은 홈페이지에 새로 만들어진 디렉토리와 파일 목록을 RSS로 주기적으로 퍼블리싱해주었다. RSS 수집 프로그램을 사용하여 사람들은 콘텐츠가 업데이트 되는 상황을 모니터링 할 수 있게 되었다.

### 4. CONCEPT P2P

#### 4.1 시스템 구조

CONCEPT P2P에서 사용자는 provider와 consumer의 두가지 역할을 가지고 P2P Network과 웹에 접근한다. 사용자의 디렉토리구조, 즐겨찾기 등을 사용하여 관심 개념 목록을 추출해서 RDF/XML포맷으로 매핑하여 개념지도도를 작성한다. 작성한 개념들 중에서 최근 업데이트 된 내용을 웹과 P2P 환경에 퍼블리싱한다. 웹에 퍼블리싱 된 개념들은 RSS 구독기를 사용하여 모니터링 할 수 있다. P2P에 퍼블리싱 된 개념지도도는 EDUTELLA 네트워크에 연결되어 다른 사용자(Peer)를 검색하는데 쓰인다.

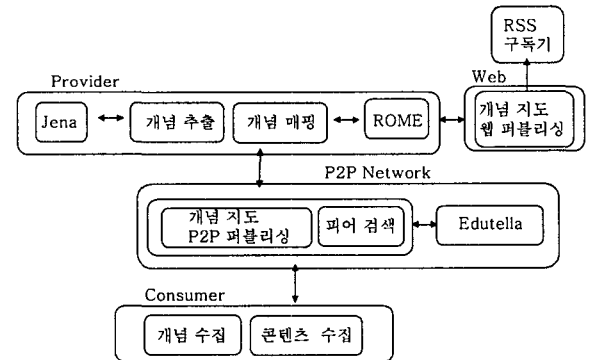


그림 2 CONCEPT P2P의 시스템 구조

밝게 표시된 모듈은 구현에서 사용한 주요 라이브러리이다. RDF 구문의 작성과 해석과 관련된 부분은 Jena [9] 파서를 사용하였고 P2P 환경에서의 동작은 EDUTELLA 컴포넌트를 사용하였다. ROME [10]은 RSS, ATOM 등의 다양한 콘텐츠 신디케이션 포맷을 지원하는 자바 API이다. 웹에 퍼블리싱 된 RSS 파일은 해당 버전의 RSS를 지원하는 RSS 구독기를 사용하여 새로 생성된 콘텐츠를 모니터링 할 수 있다.

#### 4.2 Provider 개념지도의 RDF/XML 표현

예를 들어 P2P라는 디렉토리에 Edutella, Lomster, LionShare이라는 서브 디렉토리를 만들고 자료를 수집하는 사람이 P2P라는 디렉토리를 콘텐츠 공유지로 설정하게 되면 기본적으로 다음과 같은 형식의 개념지도를 작성하게 된다.

```
<rdf:RDF
  xmlns:dcq="http://dublincore.org/2000/03/13/dcq#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <rdf:Description rdf:nodeID="A0">
    <rdf:type
  rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Bag"/>
    <rdf:_1>p2pcj.pdf</rdf:_1>
    <rdf:_2>a.pdf</rdf:_2>
    <rdf:_3>11.doc</rdf:_3>
  </rdf:Description>
  <rdf:Description
  rdf:about="http://h.ac.kr/~me/Concept/P2P/">
    <dc:creator>yepyep@ihanyang.ac.kr</dc:creator>
    <dc:title>P2P</dc:title>
    <dcq:haspart
  rdf:resource="http://h.ac.kr/~me/Concept/P2P/LionShare"/>
    <dcq:haspart
  rdf:resource="http://h.ac.kr/~me/Concept/P2P/Edutella"/>
    <dcq:haspart
  rdf:resource="http://h.ac.kr/~me/Concept/P2P/Lomster"/>
    <dcq:haspart rdf:nodeID="A0"/>
  </rdf:Description>
</rdf:RDF>
```

#### 4.3 Consumer의 개념지도 질의 QEL

데이터로그(datalog)는 매우 제한된 형태의 프롤로그와 유사한 술어 논리에 의한 질의를 표현하기 위한 언어이다. QEL은 데이터로그 시맨틱에 기반하고 있는 질의 언어이다. [11] EDUTELLA 질의 서비스를 사용하면 P2P 네트워크에서 존재하는 RDF/XML 개념지도 데이터베이스를 대상으로 검색할 수 있다. 이 때 사용되는 언어는 QEL 시맨틱과 RDF 문법을 따른다.

#### 4.4 Edutella P2P Network

P2P 네트워크에서는 기본적으로 사용자 등록과 서비스 탐색, 질의 전송, 메시지 교환이 이루어진다. 사용자를 그룹에 등록시킨 다음 네트워크 안에 존재하는 사용자 및 제공 서비스를 검색한다. 그리고 사용자는 provider로 참여하고 있는 사용자에게 질의를 보내어 자신이 원하는 콘텐츠 목록을 얻어올 수 있다.

본 시스템에서는 데이터베이스와 질의 표현이 RDF/XML의 형태로 이루어지기 때문에 P2P 네트워크상에서 콘텐츠를 공유하기 위해서는 RDF 질의 서비스를 해석할 수 있는 시스템을 검색하고 모니터링 하는 작업이 필요하다. 이러한 작업을 위해서 에듀텔라 네트워크에 사용자를 등록하고 이벤트 리스너를 구현하여 사용자

의 온라인/오프라인 상태와 EDUTELLA 네트워크에 보내진 질의 결과를 모니터링 한다.

#### 4.5 개념 지도 웹 퍼블리싱

개념지도 웹 퍼블리싱은 파일의 생성날짜를 비교하여 개념지도 중 최근에 변경된 내용을 XML 신디케이션 포맷으로 변환하는 것이다. 사용자가 자신의 홈페이지 URL을 시스템에 등록하면 FTP를 사용하여 XML 신디케이션 포맷의 개념지도를 전송하여 웹에 퍼블리싱 한다.

개념지도를 웹에 퍼블리싱을 하는 이유는 웹은 상시 접근할 수 있는 장점이 있기 때문이다. 웹에 개념지도를 퍼블리싱 함으로서 새로 등록된 콘텐츠 리스트를 모니터링 한다는 일반적인 맥락에서 웹 콘텐츠와 P2P 콘텐츠를 접근하게 된다.

#### 5. 결론 및 향후 연구과제

본 논문에서는 개념지도를 적용한 P2P 콘텐츠 공유 시스템을 제안하였다. 시스템 사용자의 개념과 관심사가 반영되는 파일 시스템 디렉토리, 즐겨찾기 등을 사용하여 관심 개념을 추출하여 개념지도를 기술하였다. 기존의 한정된 범위에서 사용되고 있던 P2P 시스템과는 달리 이 정보를 메타데이터로 사용한 P2P 콘텐츠 공유 시스템은 넓은 범위에서 활용할 수 있는 잠재력을 가지게 되었다. 또한 이 시스템은 비슷한 관심사를 가진 피어와 관계를 설정함으로써 콘텐츠를 지속적으로 수집하고 새로 생성되는 개념을 추적할 수 있도록 설계되었다.

CONCEPT P2P에서는 개인의 개념정보를 검색엔트로 사용하였다. 이 방법의 한계는 키워드와 키워드 간의 관계가 주관적인 시각에서 나타난다는 것이다. 그렇지만 분야 공통적인 동의어 의 한 온톨로지 서버가 구축되어있다면 온톨로지와 개인의 주관적인 개념을 연결 하여 콘텐츠를 배포 한다면 보다 일반적인 키워드를 사용한 콘텐츠의 검색이 가능할 것이다.

#### 참고 문헌

- [1] The Theory Underlying Concept Maps and How To Construct Them, Joseph D. Novak, Cornell University
- [2] W3C Semantic Web <http://www.w3.org/2001/sw/>
- [3] Resource Description Framework (RDF) <http://www.w3.org/RDF/>
- [4] RDF Site Summary 1.0 (RSS)
- [5] SUN, JXTA Technology Creating Connected Communities, 2004-1
- [6] JXTA, <http://www.jxta.org>
- [7] EDUTELLA <http://edutella.jxta.org>
- [8] Ben Hammersley, Content Syndication with RSS, O'REILLY
- [9] HP Jena <http://jena.sourceforge.net>
- [10] ROME, <https://rome.dev.java.net/>
- [11] RDF Query Exchange Language(QEL) - concepts, semantics, and RDF syntax
- [12] Wolfgang Nejdl, Boris Wolf, Steffen Staab, Julien Tane EDUTELLA: Searching and Annotating Resources within an RDFbased P2P Network