

모바일 무선 네트워크 환경에서 에너지 효율적인 XML 데이터 방송 기법*

박상현⁰ 최재호 이상근
고려대학교 컴퓨터학과
{newtypus⁰, redcolor25, yalphy}@korea.ac.kr

An Energy Efficient XML Data Broadcasting Method in Mobile Wireless Network

Sang-Hyun Park⁰, Jae-Ho Choi, SangKeun Lee
Dept. of Computer Science and Engineering, Korea University

요 약

본 논문에서는 모바일 무선 네트워크 상에서 효율적으로 XML 문서를 방송하기 위한 기법을 제안한다. 이를 위해 XML 문서의 구조정보를 무선 방송 데이터의 색인으로 활용하여 XML 문서를 스트림으로 재구성한다. 방송되는 XML 스트림의 색인은 XML 경로 요약물 이용하여 축약된 형태로 표현하며, 구조 정보를 XML 스트림의 색인으로 사용하기 위해 XML 문서는 구조정보와 콘텐츠 정보를 분리한다. 축약된 색인을 이용하여 각 클라이언트는 방송되는 XML 스트림에 대하여 지역적으로 질의를 수행함으로써 전체 XML 문서 중 자신이 원하는 정보만을 수신한다. 성능평가를 통해 제안하는 기법이 접근 시간 및 튜닝시간에서 많은 이득을 볼 수 있음을 보인다.

1. 서 론

무선 이동 단말 장치와 무선 통신 기술의 발달은 사용자로 하여금 원하는 정보를 언제 어디서나 접근할 수 있도록 하였다. XML(eXtensible Markup Language)[1]이 인터넷상에서 정보 교환과 데이터 표현의 표준이 됨으로써 많은 정보가 XML 형태로 사용되고 있으며, XML의 연구분야 또한 무선환경으로 확대되고 있다[2]. 그러나, 모바일 무선 컴퓨팅 환경에서는 작은 대역폭 및 잦은 끊김과 같은 제약 사항이 있으며, 기존의 on-demand 방식의 데이터 전송에 비해 무선방송(Wireless Broadcast) 기법의 경우 적은 대역폭으로도 다수의 클라이언트가 데이터를 수신할 수 있다는 장점이 있다[3]. 무선으로 데이터를 방송하는 경우 별도의 색인을 사용함으로써 클라이언트는 한 방송 주기 동안 전송되는 모든 데이터를 수신할 필요가 없다. 즉, 클라이언트는 방송 데이터의 색인을 먼저 접근하여 원하는 데이터의 도착시간을 예측함으로써 자신이 필요하지 않는 데이터의 전송 기간 동안에는 에너지 소모를 줄이는 휴지상태를 유지하고, 실제 요청한 데이터가 방송되는 동안에만 방송 채널에서 데이터를 수신하는 활성상태를 유지한다.

이러한 XML과 무선 방송 기법의 장점을 이용하여 본 논문에서는 다수의 클라이언트에게 XML 문서를 전송하는 기법을 제안한다. XML 문서를 무선 방송환경에 적합한 스트림으로 변경 시 별도의 색인을 생성하지 않고 XML 문서의 구조정보를 경로 요약 방법으로 축약하여 스트림의 색인으로 사용한다. 구조정보를 스트림의 색인으로 사용하기 위해서 XML 문서의 구조정보와 콘텐츠 정보를 분리한다. 클라이언트는 방송주기 동안 XML 문서의 구조정보를 포함하고 있는 색인을 먼저 수신하고, 색인정보와 자신이 질의를 비교하여 원하는 콘텐츠가 언제 방

송될지 확인한다. 이를 통해 클라이언트는 전체 XML 데이터들을 필요없이 필요한 데이터가 방송되는 시점에만 무선 채널에 튜닝하여 데이터를 수신할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 관련 연구를 소개하고 본 연구의 동기를 살펴본다. 3장에서는 제안하는 XML 스트리밍 기법과 질의 처리 방법에 대하여 기술한다. 4장에서는 제안하는 기법의 성능평가 결과를 분석하며, 5장에서는 향후 연구 방향에 대해 살펴보고 결론을 내린다.

2. 관련 연구

네트워크를 통한 XML 문서의 전송과 관련된 기존의 연구는 크게 세 가지 분야에서 활발하게 진행되어 왔다. 첫째는 XML 데이터의 선택적 보급에 관한 연구이다[4]. 많은 양의 XML 데이터를 사용자의 요구사항이 기술된 사용자 프로파일(profile)에 기반하여 선택적 정보 배포(Selective Dissemination of Information)를 수행하는 연구로서 XML 정보 배포와 함께 질의 처리 기법들 또한 활발하게 연구되고 있다. 둘째는 XML 스트리밍에 관한 연구로서 데이터양의 급격한 증가와 함께 실시간으로 입력되는 XML 데이터를 스트림으로 간주하여 이를 처리하는 기법이 연구되고 있다[2]. 마지막으로, 네트워크를 통한 XML 데이터의 전송에 관한 연구이다[5]. 비교적 큰 XML 데이터를 네트워크를 통해 전송하기 위해 데이터 압축이나, XML 문서내의 중복을 제거함으로써 효율적으로 XML을 전송하기 위한 연구가 진행되고 있다. 위와 같은 연구들은 XML 데이터 처리를 위한 효과적인 해결책으로 널리 인식되고 있다. 또한, 무선 모바일 데이터 관리 기법에 관한 연구는 무선 모바일 네트워크에 내재하는 여러 제약사항들(좁은 대역폭, 짧은 배터리 용량, 제약

* 본 결과물은 정보통신부의 정보통신 기초기술연구지원사업(정보통신연구진흥원)으로 수행한 연구결과입니다.

적인 컴퓨팅 능력 등)을 극복하기 위한 다양한 해결책을 제시하였다[3]. 그러나 이러한 기존의 연구에서는 XML과 모바일 무선 컴퓨팅 환경의 제약사항을 동시에 고려하지 못한다. 이를 해결하기 위해서는 모바일 무선 네트워크에 환경 적합한 XML 문서의 표현 및 전송기법이 필요하다.

3. 제안하는 XML 데이터 스트리밍 기법

무선 방송을 위한 스트림은 서버측에서 XML문서의 전처리 과정을 통해 생성한다.

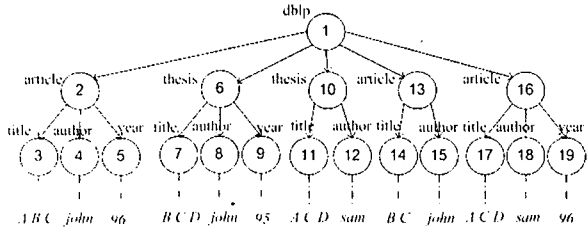


그림 1. XML 데이터 트리의 예

XML 문서는 그림1과 같이 XML 데이터 트리 형태로 표현될 수 있다. 엘리먼트 노드는 유일한 식별자를 가진 원으로 표시되었으며 점선으로 연결된 노드는 실제 콘텐츠 정보를 나타낸다. XML 데이터 트리 D 에서 노드 v 에 대한 레이블 패스(label path)를 $L(v)$ 라고 하고 각 노드를 점으로 구분한다고 정의하자. 이때 8번 노드는 루트로부터 $1 \rightarrow 6 \rightarrow 8$ 의 경로를 통해 도달할 수 있으며, 레이블 패스 $L(8) = dblp.thesis.author$ 로 나타낼 수 있다.[6] 이때 8번과 12번 노드는 동일한 레이블 패스를 가진다. 이와 같이 한 XML 문서내에서는 동일한 레이블 패스의 중복이 많이 발생하게 되며, 이러한 노드들을 그룹으로 묶어 레이블 패스의 중복을 제거함으로써 원본 XML 문서의 경로 요약본(path summary)을 만들 수 있다. 스트림 생성을 위한 전단계 처리과정으로서 서버측에서는 그림 2와 같이 XML 데이터 트리의 경로 요약본을 생성한다. 경로 요약본은 원본 XML 데이터 트리의 노드 식별자를 기준으로 정렬된 형태를 가지며, 하나의 XML 문서에 대하여 유일한 경로 요약본이 생성된다[7]. 이렇게 생성된 경로 요약본은 중복된 레이블 패스의 제거로 원본 XML 문서보다 작은 크기를 가진다. XML 문서의 경로 요약본으로부터 무선환경에 적합한 형태의 스트림으로 재구성한다.

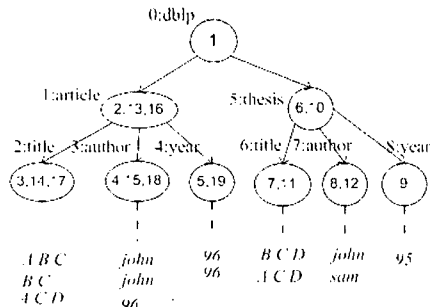


그림 2. 정렬된 path summary

XML 스트림은 그림3과 같이 경로 요약본의 넓이 우선으로 방송한다. 회색으로 나타낸 부분은 XML 데이터의 축약된 구조로서 스트림상의 색인으로 사용하게 되며 색인 다음에 실제 콘텐츠를 방송한다.

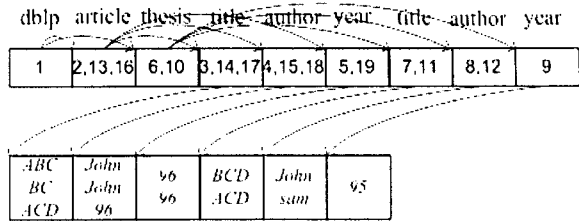


그림 3. path summary로부터 생성한 XML 스트림

이때 데이터 스트림은 그림4와 같이 색인과 콘텐츠가 분리된 형태를 가진다. 색인에서는 각 엘리먼트 노드들의 유일한 식별자인 노드ID와 해당하는 엘리먼트 이름을 가지며, 스트림내의 자신의 위치로부터 다음 자식노드까지의 거리값(바이트수)을 가진다.

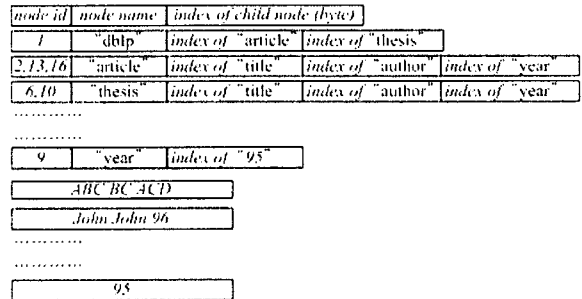


그림 4. 데이터 스트림이 가지는 정보

데이터 스트림에 대하여 클라이언트는 자신이 원하는 정보를 얻기 위해 XPath[8]형태의 질의를 사용하게 되는데, 색인 정보를 통해 자신의 질의와 일치하는 노드를 찾는다. 예를 들어 /dblp/thesis/year 와 같은 질의에 대한 답을 얻고자 할 경우 색인내의 dblp 노드를 읽으면 다음 자식노드인 thesis가 방송될 시점을 알 수 있으며, 중간에 방송되는 스트림은 듣지 않아도 되는 것이다. 또한 경로 요약을 이용한 축약된 색인을 사용하였기 때문에 원본 XML문서에서 존재하는 중복된 레이블 패스에 대한 질의 결과를 한번에 얻을 수 있다.

4. 성능 평가

제안하는 방법의 성능 평가를 위해 최근 발표된 XML 데이터 방송기법[10]에서 사용한 실험 데이터 및 실험 방법을 채택하였다. 원본 XML문서를 제안하는 기법에 맞는 파일 스트림으로 변환하였으며 실험 데이터는 [9]의 데이터를 사용하였다. 원본 XML문서와 생성된 데이터 스트림의 크기를 비교하면 표1과 같다. 원본 XML문서의 경로 요약본을 이용하여 중복된 레이블 패스를 제거함으로써 원본에 비해 많은 비율로 용량이 줄어드는 것을 알 수 있다.

표 1. 원본 XML과 스트림의 용량비교

실험 데이터	원본 XML용량 (bytes)	데이터 스트림 용량 (bytes)	비율 (스트림/원본)
SigmoidRecord	478416	245760	51.36 %
UWM	2342912	1048576	44.76 %
WSU	1654784	884736	53.47 %

각 실험 데이터에 대한 질의 내용은 다음과 같다.

표 2. 실험 데이터별 질의내용

실험 데이터 및 질의 내용
SigmoidRecord.xml
Q1 : SigmoidRecord/issue/articles/article/title
Q2 : SigmoidRecord/issue/articles/article/authors/author
UWM.xml
Q3 : root/course_listing/section_listing/instructor
Q4 : root/course_listing/section_listing/hours
WSU.xml
Q5 : root/course/days
Q6 : root/course/times

한 XML 문서에 대하여 두 가지 종류의 완전 경로 질의를 사용하여 성능을 측정하였다. 실험시 성능 평가는 전체 XML 데이터를 읽는 시간에 대하여 제안된 기법으로 XML 데이터를 읽는 시간의 비율을 측정하며 다음과 같은 식으로 정의한다[10].

$$\frac{\text{제안한 기법에 의해 측정된 시간 (ms)}}{\text{전체 데이터를 읽을때 소요되는 시간(ms)}}$$

그림 5는 각각 접근시간(Access Time)과 튜닝시간(Tuning Time)을 측정 한 결과이며, 그래프 중 실선이 본 논문에서 제안하고 있는 기법을 나타낸다. 최근에 발표된 XML 데이터 방송 기법인 S-노드[10]에 비해 접근시간 및 튜닝시간에서 나은 성능을 보여주고 있다. 전체 XML 데이터를 읽을 경우 보다 약 50% 가량의 접근시간이 소요되며, 약 15% 정도의 튜닝시간이 소요되는 것을 알 수 있다. 즉, 접근시간의 경우 전체 데이터를 읽는 경우에 비해 2배가량 빠른 시간에 원하는 데이터를 클라이언트가 얻을 수 있으며, 에너지 효율성 면에서는 6배 이상 배터리 사용시간을 늘릴 수 있다.

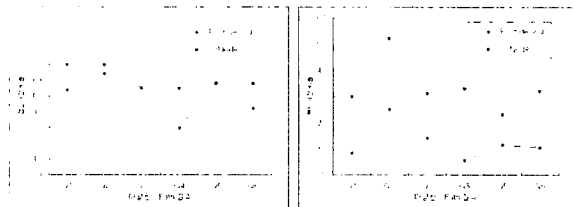


그림 5. 접근시간 및 튜닝시간 측정 결과

5. 결론 및 향후 연구

본 논문에서는 모바일 무선 네트워크 상에서 효율적으로 XML 문서를 방송하기 위한 기법을 제안하였다. 이를 위해 원본 XML문서의 경로 요약본을 이용하여 구조정보를 스트림의 축약된 색인으로 사용하였으며, 색인을 통해 스트림에서 원하는 정보만을 수신함으로써 접근시간 및 튜닝시간에서 이득을 얻을 수 있음을 실험을 통해 검증하였다. 향후 연구로는 현재의 완전 경로 질의만이 아닌 다양한 형태의 질의 지원 기법을 개발하는 것이다. 또한 실제 통신 환경에서는 일정한 크기의 패킷 단위로 데이터 전송이 이루어진다. 이러한 통신환경을 고려하여 네트워크 전송비용을 최소화 할 수 있는 XML 스트림의 패킷화 기법에 관한 연구도 이루어질 것이다. 또한, 사용자들의 데이터 요구패턴을 파악함으로써 XML 스트림의 색인과 컨텐츠 정보의 배치를 최적화시켜 접근빈도가 높은 데이터에 대한 접근시간과 튜닝시간을 현재보다 향상시킬 수 있는 연구를 진행할 것이다.

참고문헌

- [1] Extensible markup language. <http://www.w3.org/XML>
- [2] Wong E.Y.C., Chan A.T.S., Hong Va Leong, "Xstream: A middleware for streaming xml contents over wireless environments", *IEEE Transactions on Software Engineering*, 30(12):918-935, 2004.
- [3] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast Disks: Data Management for Asymmetric Communication Environments", In *Proceedings of ACM SIGMOD Conference on Management of Data*, pages 199-210, 1995.
- [4] M. Altinel and M. J. Franklin, "Efficient filtering of xml documents for selective dissemination of information", In *Proceedings of International Conference on Very Large Data Bases*, pages 53-64, 2000.
- [5] Y. Diao, S. Rizvi, and M. J. Franklin, "Towards an internet-scale XML dissemination service", In *Proceedings of International Conference on Very Large Data Bases*, pages 612-623, 2004.
- [6] Qinghua Zou, Shaorong Liu, Wesley W. Chu, In *Proceedings of ACM International Workshop on Web Information and Data Management*, "CTree: A Compact Tree for Indexing XML Data", pages 39-46, 2004
- [7] R. Goldman and J. Widom, "Dataguides: Enabling query formulation and optimization in semistructured databases", In *Proceedings of International Conference on Very Large Data Bases*, pages 436-445, 1997.
- [8] XML Path language. <http://www.w3.org/TR/XPath>
- [9] XML Data Repository. <http://www.cs.washington.edu/research/xmldataset/>
- [10] Chung Soo Kim, Yon Dohn Chung, "An XML Data Streaming Method for Wireless Broadcast in Mobile Computing Environments", In *Proceedings of Korean Database Conference*, pages 79-85, 2005.