

무선 인터넷 플랫폼을 위한 XML 기반 Contents Generator System

윤수진* , 신혜원**
고려대학교 컴퓨터과학기술대학원
e-mail : bani23@korea.ac.kr

System of Contents Generator in XML for Wireless Internet Platform

Soo-Jean Yoon* , Hae-Won Shin**
Graduate School of Computer Science & Technology, Korea University

요 약

최근 들어 무선인터넷 단말기로 보다 효율적으로 인터넷에 접근하여 정보를 제공하고자 하는 연구가 활발히 진행되고 있다. 사용자는 유선 인터넷으로 제공받는 콘텐츠를 휴대가 간편한 무선 이동통신으로 제공받고자 하는 욕구가 점점 늘어나고 있고 이러한 원인으로 다양한 무선 콘텐츠의 제공이 요구된다. 그러나 다양한 무선 단말기의 특성에 따라 콘텐츠를 새롭게 개발하는데는 개발비용, 기간이 많이 들게 되므로 기존의 웹 콘텐츠를 활용하여 무선환경에 맞는 문서변환 기술 연구의 필요성이 대두되고 있다. 본 논문에서는 유선 인터넷상의 콘텐츠를 무선 환경으로 손쉽게 변환할 수 있는 XML(eXensible Markup Language)기술을 활용한 변환기(Generator)구조를 설계하고 이와 관련된 XML Generator, RuleSet Management, Clipping Toolkit 등의 기능을 정의하여 보다 효율적으로 무선 콘텐츠로 변환하는 기술을 소개한다.

1. 서론

정보통신기술 업계는 최근 무선 인터넷이라는 화두로 뜨겁게 달구어지고 있다. 전세계의 휴대 단말기 가입자가 2000년부터 급격히 늘어 2005년에는 무선 인터넷의 시장규모가 연평균 11% 이상 성장할 것으로 예상되고 있다[1]. 이러한 무선 인터넷에 대한 사용자의 증가와 기대 요구에 부응하여 무선 인터넷 사업자들은 유선 인터넷과 같은 다양한 콘텐츠 제공이라는 과제를 안고 그 연구가 활발히 진행되고 있다. 그러나 무선 콘텐츠는 이동통신 사업자의 제공 방식에 따라 각각 새롭게 구축하여야 하는 한계를 가지고, 유선 콘텐츠 사업자가 무선 환경으로 서비스를 제공하고자 할 경우 개발 기간과 비용에 문제점을 갖게 된다. 본 논문에서는 기존에 서비스되고 있는 유선 인터넷 콘텐츠를 무선 환경의 다양한 디바이스에 구애받지 않고 빠르고 간편하게 변환할 수 있는 시스템을 설계하였다. 즉, 기존에 제공하고 있는 유선 인터넷 콘텐츠 및 서비스를 재구성하지 않고 XML 문서 규약과 구조,

내용, 표현을 자유롭게 처리할 수 있는 고급 표준 기술인 XSLT 와 XPATH 이용하여 재사용성을 극대화하는 방안으로, 기존의 콘텐츠(HTML) 유형을 템플릿화(XSL, DOM)하여 구성, 관계형 DataBase 로 저장해 두고 무선 단말기(휴대폰, PDA 등)로 부터 페이지 요청이 발생하면 웹 페이지의 Document 와 별도 보관하고 있는 템플릿을 결합하여 Target Page 로 변환하여 서비스를 제공하는 기술을 설계한다.

2. 관련 연구

2.1 무선 인터넷 동향

2004년 5월말 현재, 국내 이동통신서비스 가입자는 총 3600만여 명으로 전체 인구대비 보급율(penetration)은 약 75%에 육박하고 있다. 이러한 무선 인터넷 수요의 급격한 증가로 인해 다양한 형태의 무선 인터넷 서비스가 제공되고 있고, 최근 무선 인터넷 서비스의 휴대성, 무선랜 카드의 서비스 대역폭이라는 장점을 결합시킨 WiBro(Wireless-Broadband) 서비스가 도입되어 그 시장은 점점 확대될 것으로 예상된다[2].

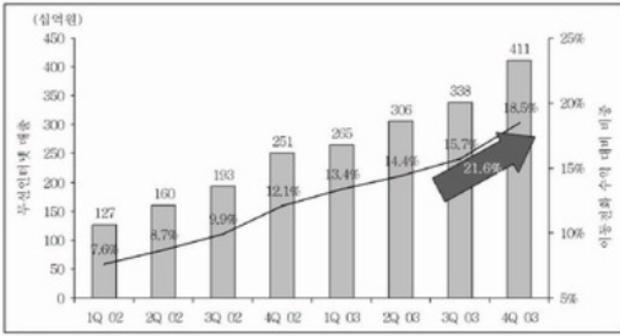


그림 2.1 무선인터넷 매출 변화 추이 (SKT기준) <자료>SKT투자자설명자료(2004.2)

국내에서는 지금까지 무선 인터넷 서비스의 한계로 인해 유선 콘텐츠와 유사한 무선 포털 서비스 등이 활성화되지 못하였으나 다양한 콘텐츠를 기반으로 한 일본 NTT-DoCoMo의 성공한 사례처럼 앞으로의 수요 시장은 매우 밝을 전망이다[3].

2.2 XML 기술을 활용한 콘텐츠 변환

지금까지 연구되어 온 콘텐츠 변환 기술 중 하나로 HTML Filter 기능을 통한 맵핑 방식이 있으나 이는 맵핑 정보에 대한 지속적인 업그레이드로 인하여 유지보수가 어렵고 정확성이 떨어져서 콘텐츠의 손실을 가져올 수 있는 단점이 있다[4]. 이후에 인터넷 상의 데이터 표현과 교환을 위한 새로운 표준인 XML(eXtensible Markup Language)을 중간 표준 포맷으로 이용하여 웹에서 제공되는 HTML 문서를 Well-Formed XML 문서로 변환, 스타일시트를 적용하여 무선 단말기에 응답처리 하는 방법이 있다. 이 방법은 서비스 제공자가 변환할 콘텐츠를 미리 분석하여 스타일시트를 직접 편집해야 하는 복잡한 과정을 거쳐야 하므로 유지보수의 문제는 해결되나 변환할 문서의 구조를 잘 파악하고 있어야 한다는 한계점이 있다[5].

2.3 DOM(Document Object Model)

DOM은 XML 문서를 조작하기 위한 API로서, 다양한 환경과 애플리케이션에서 사용할 수 있는 표준적인 프로그래밍 인터페이스를 제공한다. DOM은 문서 전체를 파싱 처리한 후 문서를 위한 문서 트리를 구성한다.

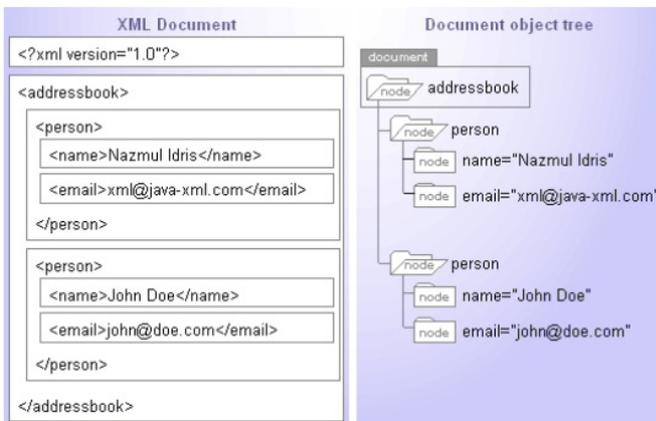


그림 2.2 DOM Tree

DOM은 메모리에 의존적이어서 XML 문서의 크기가 크면 문제가 될 수 있으나 노드의 생성, 삭제 등 데이터의 재가공이 용이하므로 매우 유용한 기술이다[6].

3. 시스템 설계 및 구현

다양한 형태의 콘텐츠를 무선 인터넷으로 서비스 받을 수 있도록 하는 Server 및 Client 측 변환기 (Generator) 시스템을 제시하고자 한다.

3.1 제안 시스템의 구조

본 변환 시스템의 핵심 기능은 다음 그림 3.1에서와 같이 Request를 받아 Response를 실행하는 Server Side와 System Rule을 편집하는 Clipping ToolKit으로 볼 수 있다.

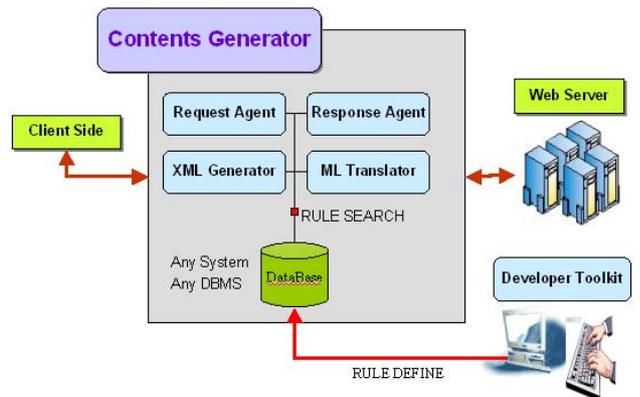


그림 3.1 전체 시스템 구조도

3.2 Contents Generator Server Demon

콘텐츠의 HTML Source를 로드하여 XML로 변환, DB에서 저장된 XSL을 가져와서 TargetML을 생성하여 서비스를 제공한다. Contents Generator Server Demon은 Browser(Any Phone, Web)가 실제 존재하는 Web Page를 요청하면 G.conf에 의하여 실제 URL을 획득하고 접속 후 요청 페이지의 응답 내용을 가져와서 해당문서를 Well-Formed XML로 변환시킨다. DB에 저장되어 있는 XSL(Clipping Tool로 만들어진)과 Well-Formed XML을 조합하여 나온 Target ML(wml, html, chtml, hdml, ...) 데이터 결과를 다시 요청 Browser에 뿌려주는 역할을 한다.



그림 3.2 G.Conf 파일 구조

3.3 Clipping ToolKit

인터넷에서 제공되는 콘텐츠의 페이지 소스(HTML) 중 원하는 콘텐츠를 Clipping 하는 기능을 제공한다.

Clipping 된 페이지의 템플릿은 XSL 화 되어 확인될 수 있고, 작업을 완료하면 페이지 구성 RuleSet 을 DB 에 저장할 수 있다.

- URL 을 입력받아 Navigating 한다(Browser 기능)
- HTML 문서를 분석하여 프레임구조를 Tree 형태로 출력할 수 있다.
- DB Accessing : HTTP Response 를 분석하여 Header 부분의 정보와 Body 부분의 Frame Tree 에 관한 정보들을 DB 에 저장
- Well-Formed HTML 문서 생성
- DOM Tree : Well-Formed HTML 문서를 DOM Parsing 을 통해 Document Node 생성. Document Node 를 이용하여 Tree Control 에 Node Type, Name, Type 으로 구분표시, 편집 가능
- Edit Tree : DOM Tree 내용을 수정하여 새로운 Tree 형태로 편집 가능. Edit Tree 를 통하여 XSL 문서 생성
- XSL Creating : Edit Tree 를 XSL 문서로 변환. XSL 을 Text Mode 로 편집 가능
- Target ML 생성 : Well-Formed HTML 과 XSL 을 결합하여 Target ML 을 생성

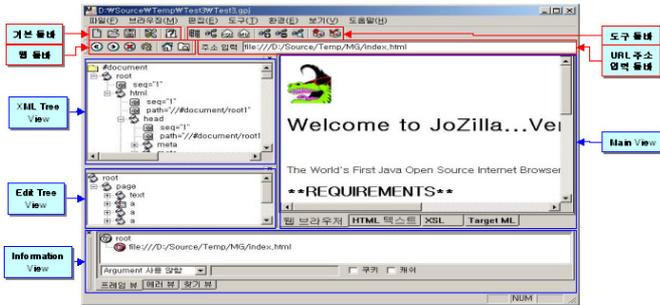


그림 3.3 제안 시스템의 Client Toolkit

3.4 RuleSet 정의

컨텐츠 변환 구성 규칙에 대한 정보를 별도의 관계형 DabaBase 에 저장하고 Response Agent 는 해당 컨텐츠의 RuleSet 정보를 가지고 서비스를 제공할 수 있도록 한다.

3.4.1 DB 구성

- C_URL : 변환할 컨텐츠의 URL 정보(인수값 제외)
- C_COOKIE : 웹 페이지의 COOKIE 사용여부
- C_ARG_LIST : URL 의 ?뒤 인수 정보
- C_FRAME : 웹 페이지의 FRAME 저장
- C_FRAME_SRC : 웹 페이지의 FRAME SRC 저장
- C_PRIORITY : 프레임의 변환 우선순위 저장
- C_BROWSER_TYPE : XSL 생성자의 브라우저 형식

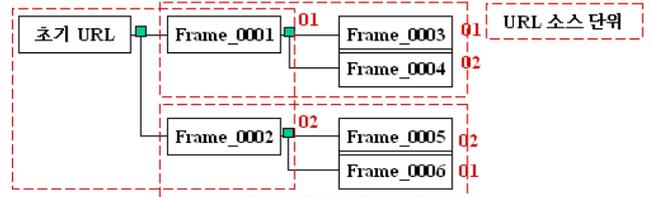
저장(WML, ME,...)

- C_XSL : XSL 문서 저장
- C_CACHE : 웹 브라우저의 CACHE 사용여부
- C_COMMENT : 주석

3.4.2 Frame 처리 및 URL 변경 규칙

DB 에서 우선순위 정보를 읽어와서 우선순위가 높은 Frame(Form)의 소스문서를 조회해 온다. 우선 순위는 같은 URL 소스에서 파생된 폼들 만이 비교 대상이 되도록 지정하여 DB 에 저장한다.

□ 샘플 폼 구성 및 우선 순위 형태



□ 처리

```
SELECT C_PRIORITY, C_FRAME
INTO 변수
FROM tb_xsl
WHERE C_URL = URL 변수
AND C_BROWSER_TYPE = 브라우저 종류
AND C_ARG_LIST = 매개변수(or 매개변수타입);
```

그림 3.4 우선 순위 규칙

3.4.3 XSL 생성규칙

- WEB Server 에 HTML 문서요청
- WEB Server 에서 받아 온 HTML 문서를 Well-Formed XML 로 변환
- XML 문서 항목 중 무선 인터넷 서비스하고자 하는 부분을 선택하여 XSL 문서로 DB 저장

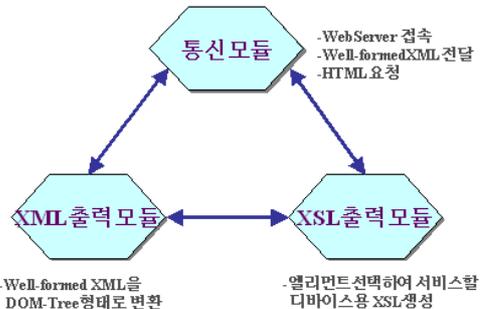


그림 3.5 XSL 변환 구성도

3.5 개발환경 및 평가

제안 시스템은 기존 컨텐츠 변환 설계와 같이 스크립트를 직접 편집하는 것이 아니라 GUI 환경의 ToolKit 을 제공하여 컨텐츠 변환 규칙을 정의하고 PreView 할 수 있어 작업자가 보다 빠른 시간 내에 쉽고 간편하게 작업을 처리할 수 있다. 본 시스템의 설계를 통하여 다음과 같은 기대효과를 얻을 수 있다.

- Client 환경에서 무선 인터넷 서비스 변환 정의를 개발 / 저장 / 관리가 가능하다.
- 순수 JAVA 로 구현하여 어느 Platform 이나 디바이스 종류에 독립적인 서비스가 가능하다.

- 정보의 저장 형태를 일반적인 RDBMS 을 사용하여 특정 DBMS 에 종속되지 않는다.
- HTML, XML 을 Tree 형식화, 또한 Tree 와 문서와의 연동을 통하여 사용자의 편의를 도모한다.
- 각종 추가기능을 이용하여 HTML 변환의 한계를 극복한다.
- 비 전문가도 손쉽게 적응할 수 있는 User Interface 를 구축한다.
- 기 존재하는 기능/내용이외에 추가로 기능/내용 추가할 수 있는 유연성을 가진다.
- 무선 인터넷 서비스 개발 기간이 단축되어 개발 비용 절감 및 생산성이 향상된다.

제안 시스템의 개발환경은 아래 표와 같다.

| 환경 | 구현 기술 | |
|------------------|-----------------------|----------------------|
| OS | Unix Sun, Windows2000 | |
| Program Language | Server | JAVA(JDK1.4) |
| | Client | Microsoft Visual C++ |
| DBMS | Oracle8i | |

표3.1 개발 환경

4. 결론 및 향후과제

국내 무선 인터넷 서비스는 가입자의 수적인 증가 및 서비스의 다양성에 있어서 세계에서 우위를 점하고 있으나, 표준화의 미비로 인하여 콘텐츠 사업자는 이동통신 사업자별 플랫폼에 맞는 콘텐츠를 중복 투자하여 개발하는 등 무선 인터넷 사업 확장에 어려움이 따르고 있다[10]. 또한 무선 인터넷의 발달로 인하여 기존의 유선 통신 사업자들도 무선 통신으로 그 영역을 확대하는 데 있어 다양한 플랫폼별 콘텐츠 제작에 개발비용 및 생산성에 많은 어려움을 안고 있다. 이에 본 논문에서는 유선 콘텐츠를 변환하는데 있어 디바이스 종류에 독립적이고 유연성 있는 시스템을 설계하여 개발 생산성을 향상시키는 효과를 얻을 수 있게 하였다.

향후 마크업 언어 단위의 변환 규칙이 아닌, XHTML 의 모듈성을 이용한 모듈 단위의 변환 규칙을 지원할 수 있도록 하는 연구와 무선 인터넷을 통한 전자금융 및 전자상거래 등의 작업을 지원하기 위한 기본적인 보안에 대한 방안 등의 연구 등을 보완해 나갈 예정이다.

참고문헌

[1] 여재현, “wibro 서비스”, IT 산업전망컨퍼런스 2005.
 [2] 이홍재, “이동통신 산업의 최근 동향과 전망”, KISDI 이슈리포트, 2004. 07.
 [3] 성균관대학교 IT 경영연구센터, “국내 무선인터넷시장의 현황과 기업의 무선인터넷 활성화방안에 관한 연구”, 정보통신연구진흥원 보고서, 2003. 07.
 [4] 이승진, “확장성 있는 웹 서비스를 위한 무선 응용 프로토콜 기반의 HTML Filter 구현”, 정보과학회

봄학술발표논문집(A), pp.391-393, 2001.
 [5] 김미영, “무선 인터넷 서비스를 위한 XML 기반 콘텐츠 변환기 설계”, 한국정보처리학회 춘계학술발표대회 논문집, 2003.
 [6] 김태현, “XML 소프트웨어 개발에서 DOM활용 및 DTD 모호성 제거를 위한 디자인 패턴”, 연세대학교 석사논문, 2000.
 [7] “XSL Transformation(XSLT) Version 1.0”, W3C Recommendation, 16, November 1999.
<http://www.w3.org/TR/XSLT>
 [8] WAP Forum, “Wireless Markup Language version 2 Specification,” June 2000.
<http://www.openmobilealliance.org>
 [9] The W3C’s XML Web page, 1998,
<http://www.w3c.org/XML>
 [10] 한국정보통신기술협회, “무선인터넷 표준화 포럼 운영”, 2003년도 정보통신표준화전략포럼 최종연구보고서, 2003. 12.