

WAP기반 무선 콘텐츠 변환기 System의 설계 및 구현

이정은^o, 최철웅, 장지산, 김성찬, 신동일, 신동규
세종대학교 컴퓨터공학과
e-mail : {Jelee, cwchoi, Jsjang, kimschan, shindk,
dshin}@gce.sejong.ac.kr

A Design and Implementation of Wireless Contents Converting System based on Wireless Application Protocols

Jungeun Lee, JiSan Jang, Minsoo Kim, Sungchan Kim
Dongkyoo Shin, Dongil Shin
*Dept of Computer Engineering, Sejong University

요약

본 논문에서는 기존의 HTML로 기술된 웹 페이지를 그대로 유지하면서 휴대 단말기를 이용하여 웹 페이지에 접근하려고 할 때, 휴대 단말기에서 웹 페이지를 기술하는 언어인 WML로 변환하여 사용자에게 서비스하는 시스템을 개발하며, 좀 더 고질의 서비스를 위해 웹 에이전트를 이용하여 사용자의 기호에 맞는 데이터를 웹 상에서 추출하여 WML 문서로 전환해주는 무선 콘텐츠 변환기 System을 설계 그리고 구현한다.

1. 서론

1990년 이후 유선망을 중심으로 한 인터넷의 비약적인 발전으로 사용자는 언제 어디에서나 개인용 컴퓨터와 네트워크 연결이 있는 곳이면 자신이 원하는 정보를 손쉽게 취득할 수 있게 되었다. 현재 이동통신망에 가입하여 휴대용 전화기나 PDA같은 휴대용 무선장비를 이용하는 사용자의 수는 국내에서만 2000만 명을 넘었으며 앞으로 개인용 컴퓨터 사용자보다 더 많은 수의 사용자가 이를 이용하게 될 것으로 보인다. 이러한 휴대용 무선장비를 통한 웹 콘텐츠 이용자의 확대는 데이터통신의 급격한 증가를 가져올 것이고, 엄청난 국내외적인 수요를 불러일으킬 전망이다.

WAP[3]의 출현은 이러한 일련의 사용자의 요구를 충족시켜주고 있으며, 인터넷상에서 WAP을 이용한 기술들이 개발되고, 이용되고 있다. 그러나 기존의 PC에서 웹 페이지를 기술하는 언어인 HTML[1,5]은 PC와 같은 대형화면에서 정보를 표현하게끔 만들어

졌기 때문에 휴대용 단말기의 낮은 대역폭과 작은 디스플레이 화면, 그리고 낮은 전력과 같은 PC에서는 문제가 되지 않았던 제약 사항들을 가지고 있어 여기에 맞는 웹 페이지 기술언어를 따로 만들지 않으면 안되었다. 이러한 요구사항에 따라 핸드폰 상에서의 웹 페이지 기술언어인 WML[2]이 만들어졌다. 이런 기술적인 문제로 인해 웹 상에서의 똑같은 정보를 상이한 입출력 인터페이스로 인해 다른 언어로 기술해야한다는 단점이 발생되었다. 이로 인해 같은 정보를 서비스하기 위해 두 개의 사이트를 운영해야 하는 불합리적인 면과 웹 페이지의 개수를 생각해볼 때, 거대한 액수의 개발비용과 유지비용, 시간, 작업이 들어가게 된다.

이에 본 논문에서는 기존의 HTML로 기술된 웹 페이지를 그대로 유지하면서 휴대 단말기를 이용하여 웹 페이지에 접근하려고 할 때, 휴대 단말기에서 웹 페이지를 기술하는 언어인 WML로 HTML 문서를 변환하여 사용자에게 서비스하는 시스템을 제안

하여 설계 또 구현한다.

2. 관련 연구

본 장에서는 모바일 환경에서 사용되는 콘텐츠 구현언어의 현재까지의 동향에 대해 분석해보고, 이중 가장 널리 사용되는 WML 언어로의 필터링을 위해 HTML과 XML[4,6] 그리고 WML 언어 각각에 대해 분석해본다.

2.1 HTML 문서에 대한 분석

HTML[14,21] 문서는 시작태그(start tag), 콘텐츠(content), 끝태그(end tag)와 같이 3부분의 기본 형태를 가지며, 엘리먼트(element)들이 트리형태로 구성된다. 그리고 HTML 스펙(Specification)은 2가지 형태의 엘리먼트 타입을 정의하고 있다. 먼저 2개의 태그 사이에 콘텐츠가 들어있는 non-empty element와 단지 한 개의 태그로만 정의되고 콘텐츠를 가지고 있지 않는 empty-element이다.

HTML문서의 특징은 다음과 같다.

- HTML은 사용자에게 의해 확장될 수 없으며, W3C에 의해서만 새로운 엘리먼트를 추가함으로써 확장가능
- HTML은 사용자가 표현하고자 하는 문서의 의미를 표현하는데 있어서 충분한 수단을 제공하지 못함.
- HTML은 empty element를 찾기 위한 규칙을 제공하지 않음.

2.2 XML 문서에 대한 분석

Extensible Markup Language(XML)[11,17]은 W3C(World Wide Web Consortium)에 의해 정의되었으며 웹을 위해 디자인된 SGML에 기반을 둔 마크업 언어이다. HTML 처럼 XML 역시 태그 기반의 언어이며, 엘리먼트들을 트리 형태로 표현함으로써 문서를 기술한다. XML 스펙은 XML 문서를 valid 문서와 well-formed 문서의 2가지 형태를 제공한다. valid 문서는 문서에 대한 DTD(Document Type Declaration)를 정의함으로써, 문서의 파싱 과정에서 이 XML 문서가 DTD에 의거하여 유효한 문서인지를 검사하고, 유효한 문서일 때, 문서를 처리하는 것을 말하며, Well-formed 문서는 문서의 DTD와는 상관없이 문서의 파싱 과정에서 문서가 XML 스펙에서 지정하는 일정한 규칙을 - 시작태그

와 끝태그가 반드시 기술 - 따른다면 유효한 문서로 인정하여 문서를 처리하는 것을 말한다.

2.3 WML문서에 대한 분석

WML[1,18]은 태그에 기초한 문서언어이며, XML에 기반을 둔 일반적인 마크업 언어이다. WML은 HDML 2.0의 일부분에 기초를 두고 있다. WML은 HDML로부터 적용된 몇 가지 요소를 바꾸고 HTML과 비슷한 요소를 가진 새로운 일부 요소를 추가하였다. WML에서도 카드(card)와 데크(deck)라는 개념을 사용한다. 이것은 응용프로그램이 다중 카드로 만들어진 문서(일반적으로 데크라고 말한다)를 나타낼 수 있게 설계되어 있다. 또한 WML은 XML 문서 형식으로서 나타내어지며, 전화나 다른 무선이동터미널과 같은 제한된 능력을 가진 장치 상에서 프리젠테이션과 사용자의 상호작용을 나타내는데 최적화 되어있다.

3. WAP 필터링 시스템의 설계 및 구현

본 장에서는 먼저 무선 콘텐츠 변환기 System의 전체 구조에 대하여 알아보고, 핵심 모듈인 HTML2WML engine에 대하여 기술한다. 그리고 무선 단말기에 효과적으로 문서를 표현하기 위해 본 구현에서 적용한 문서구조 방식에 대해 기술하였으며, 변환된 문서를 효과적으로 표현하기 위한 사용자 인터페이스 및 전체 구현환경에 대해 기술한다.

3.1 시스템의 내부구조

무선 콘텐츠 변환기는 [그림 1]과 같은 구조로 구성된다.

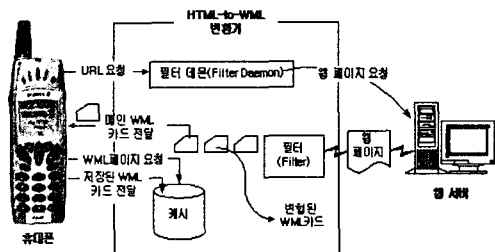


그림 2 무선 콘텐츠 변환기 전체구조

Filter 데몬을 통해 들어온 웹페이지는 소스를 파싱하고, 파싱을 통해 얻어진 document 객체를 해당 문서를 처리할 수 있는 클래스 즉,

HTMLTransform()과 HDMLTransform()으로 보내어 문서를 WML 문서로 변형하게 된다. HTML 문서는 기본적으로 90개 이상의 태그로 구성되며, 각각의 태그의 사용에도 예외상황이 많아 완벽하게 HTML에서 제공하는 문서의 특징을 WML로 재생활 수 없다. 그러므로 WML로 표현 가능한 HTML의 태그를 찾아내기 위해 HTML 문서의 재구성이 필요하며 이렇게 변형 가능한 HTML 문서를 WML로 전환한다. 전환 시 태그의 매핑을 문서의 의미적 측면을 고려하여 WML로 전환하였다.

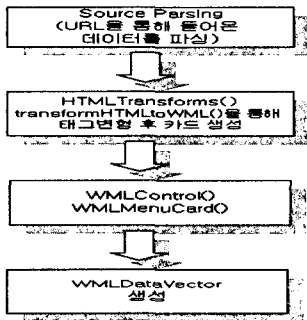


그림 3 내부 구조도

HTML에서 제공되는 태그를 아래의 논리적 요소로 분류하고 해당 태그를 아래의 [표 1]과 같이 WML 태그로 변형한다.

논리적 요소	해당 Element
Head	<head>, <access>, <meta>, <setvar>, <template>
Text	, <big>, , <i>, <small>, , <u>,
Select	<select>, <option>, <optgroup>, <fieldset>
Link	<a>, <anchor>
Card Action	<do>, <go>, <prev>, <noop>, <onevent>, <postfield>, <refresh>, <timer>
Input	<input>
Image	
Comment	<!--, -->
Table	<table>, <td>, <tr>
Paragraph	<p>

표 1 논리적 요소에 따른 WML 태그 분류

이렇게 변형된 문서는 각각의 카드로 구분되어 저장되며, [그림 2]에서 설계된 것 처럼 저장된 카드는 wmlControl()을 통해 1200 바이트로 잘려져서 카드를 구성하고, wmlMenuPage()를 통해 메뉴를 구

성한다. 페이지 구성 원칙은 첫 카드는 문서의 타이틀을 나타내며, 두 번째 카드는 문서의 메뉴 카드를 구성하여 각각의 메뉴를 통해 세부 내용을 볼 수 있는 형태의 기본 형태를 가지도록 구성된다. 웹 문서를 무선 단말기에 사용되는 WML문서로 전환할 때, 문서가 사용자에게 의미있는 문서가 되기 위해서는 무선 단말기에서는 표현되는 스크린의 크기나 나타낼 수 있는 글자수의 제한이 상당히 크게 작용함으로 이에 대한 충분한 고려가 필수적이다.[7]

이러한 과정에서 사용되는 보다 자세한 구조를 살펴본다.

■ 재구성

입력으로 들어오는 HTML문서는 먼저 openxml1.2에서 제공하는 HTML 파서를 통해 파싱된 후, 매핑 테이블을 통해 WML문서로 변환가능한 HTML문서로 다시 저작

■ Filtering & Transforming

위의 웹문서의 재구성과정을 거친 HTML문서는 xml4j3.1에서 제공하는 xml 파서를 통해서 각각의 태그와 내용을 분리한 후, 재귀적으로 동작하면서 각각의 태그를 WML태그로 변형

■ Source Generating

바뀌어진 문서는 Source Generating과정을 통해 WML문서의 파일 크기 제한규격에 맞추어 여러개의 WML 파일로 나누어지게 된다.

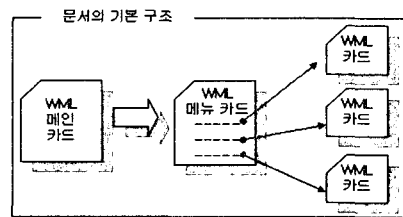


그림 4 문서 변환 구조

3.2 사용자 인터페이스

본 논문에서 구현된 사용자 인터페이스는 두 가지로 구현되었다. 하나는 WAP 필터링 시스템을 독립 실행형 자바 애플리케이션으로 구현하였으며, 또 하나는 WAP 게이트웨이와의 정합을 위해 자바 서버 데몬을 통해 WAP 필터링 시스템을 구현했다. 아래 [그림 4]는 JAVA 애플리케이션으로 제작된 HTML2WML Converter의 인터페이스이다. URL 입력란에 접근하고자 하는 사이트의 URL을 입력한 후, 버튼을 클릭하면 HTML Source 창에 해당

