

웹 콘텐츠의 구조 분석을 이용한 무선 단말기와 RSS로의 실시간 적응 변환 서비스

류 동엽*, 한승현*, 임영환**

Realtime Adaptation Transcoding Service, supporting wireless mobile devices and RSS by using Web Structure Analysis

Ryu Dong-Yeop*, Han Seung-Hyun*, Lim Young-Whan**

요약

초고속 무선 인터넷 서비스기술의 발전으로 많은 사용자들이 이동 통신 단말기를 이용한 무선 인터넷서비스를 사용하고 있다. 하지만 기존 HTML 웹 페이지의 경우 불필요한 광고와 큰 해상도의 모니터, 다양한 멀티미디어 데이터, 그리고 빠른 통신속도를 염두하고 제작된 경우가 일반적이기 때문에 무선 인터넷으로 이용하기 위한 제약이나 불편함이 많다. 따라서 본 연구에서는 기존의 웹 사이트 중 관심 영역을 RSS 채널 개념으로 재활용해서 무선 인터넷 페이지를 쉽게 생성하는 방법에 대해서 기술한다. 웹 페이지 관리자는 쉽고 빠르게 무선 페이지를 생성 할 수 있으며 무선 인터넷 사용자는 자신이 원하는 정보를 쉽고 빠르게 구할 수 있다는 장점을 가지게 된다.

Abstract

As high speed Internet service technology develops, many people are accessing wireless internet service by using mobile devices. However, an HTML web page for PC is not very compatible with wireless internet because of its unnecessary advertisements, high resolution, and various multimedia data. This is because the main focus when creating it was on high speed. As this research indicates with the RSS channel, converted from the main part of the HTML web page, which users show interest in, a wireless compatible internet page can easily be created. Consequently, web administrators could create a wireless page easily and quickly, and wireless users would be able to find information in the same manner.

▶ Keyword : RSS(RSS), 무선통신(Wireless), XML(XML), 변환(TransCoding)

• 제1저자 : 류동엽

• 접수일 : 2006.07.28, 심사일 : 2006년 8월 17일, 심사완료일 : 2006.09.20

* 송실대학교 컴퓨터 학과 ** 송실대학교 미디어 학과

※ 이 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음.(KRF-2004-005-D00198)

1. 서론

대부분의 인터넷 사용자 들은 여러 사이트를 방문해서 자신이 필요한 정보를 수집하고 있으며 최신 정보를 구하기 위해서는 다시 사이트를 방문하는 경우가 일반적이다. 하지만 이런 정보 수집방식은 비효율적이며 비교적 많은 시간을 투자해야 한다. 따라서 채널 방식의 다이나믹 콘텐츠 서비스가 무선인터넷 서비스의 주류로 떠오르고 있으며 이 서비스를 이용할 경우 사용자가 별도의 정보 접근과정이 없이도 사용자가 사전에 설정한 관심 분야의 뉴스와 상품 정보를 맞춤형으로 제공하는 방식으로 무선인터넷 서비스 패턴의 새로운 진화 모델로 주목 받고 있다[1]. 이러한 다이나믹 콘텐츠 서비스를 이용한 정보 수집방식의 개념이RSS(Really simple syndication, RDF Site Summary)와 같은 맞춤형 정보 배달 방식이다. RSS의 기본 목적은 유명 신문사의 기사를 손쉽게 제공하기 위해서였으며 성공적인 XML(Extensible Markup Language) 모델로 평가되며 현재 계속 발전하면 표준화 과정에 있다[2].

RSS를 이용하면 특정 사이트를 직접 방문하지 않고서도 RSS Reader와 같은 프로그램을 이용하여 사이트 업데이트 유무를 쉽게 확인할 수가 있으며 사용자가 원하는 정보만 쉽고 빠르게 수집할 수 있다는 장점을 가지게 된다. 이런 RSS의 장점을 무선인터넷 서비스에 접목시킨다면 무선인터넷 사용의 편리성도 증대시킬 수 있다는 장점을 가지게 된다.

본 연구의 시작은 어떻게 기존의 웹 페이지를 재활용해서 무선 인터넷 서비스를 쉽게 할 수 있는가에서 출발 했다. 하지만 기존의 웹 페이지를 무선 인터넷으로 이용하기 위해서는 사용 브라우저, 통신속도, 화면 크기, 하드웨어 특성, 지원 미디어 종류에 따른 제약이 많이 존재한다[4]. 일반적인 무선페이지 생성 방법은 다양한 무선 단말기 및 브라우저에 적합하도록 무선 인터넷 서비스를 위한 웹 사이트를 별도로 구축하는 방법과 기존 유선 웹 사이트의 콘텐츠를 변환해서 무선 인터넷 서비스로 생성하는 방법이 있다[5][6]. 별도의 무선 사이트를 구축하는 방식은 시간과 비용소모가 많고 기존 사이트와의 실시간 연동성이 어렵다는 문제를 갖는다. 변환 생성하는 방식에는 수동생성 방법과 자동생성 방법이 있으며 각 방법은 다음과 같은 특징을 갖는다. 우선 수동 생성의 경우 사이트 관리자가 수작업으로 무선 페이지를 생성할 때 최적화된 페이지를 생성 할수는 있지만 모든 단말기 환경에 맞는 무선 페이지를 생성하기 위해서는 많은 시간과 비용이 소모된다. 그리고 자동 변환을 이용하는 무선페이지 생성

의 경우 손쉽게 페이지를 생성 할 수는 있지만 생성 된 페이지의 질이 떨어지며 최적화가 어렵다는 단점을 가진다 [7][8]. 웹 콘텐츠를 변환하는 기존 연구의 경우 몇몇 단점을 가지고 있다. 첫째, HTML을 대상으로 하기 때문에 데이터 연동 관점에서 문제가 생긴다. 즉, 원본 웹 콘텐츠가 변경되면 무선페이지로 반영하기 위해서는 무선페이지를 재 생성해야 한다. 이것은 많은 비용과 시간이 소모되며 실시간 정보전달을 어렵게 한다. 둘째, 특성이 다른 모든 단말기에 대한 고려가 부족하다. 또한 하나의 웹 콘텐츠를 무선서비스 하기 위해서는 다양한 단말기를 위한 무선 페이지를 생성해야 한다는 단점을 가지게 된다. 셋째, 무선 페이지로 변환된 소스는 무선페이지만을 위해서 사용된다. 이는 하나의 소스를 다양하게 이용하지 못한다는 의미이며 변환된 무선페이지를 다양하게 사용할 수 있는 방법이 있어야 한다. 넷째, 사용자 특성과 취향을 고려하지 않는다. 따라서 하나의 웹 콘텐츠를 무선 변환할 경우 수십 개의 무선페이지가 생성되는데 이중 자신이 원하는 정보를 보기 위해서는 수십 개의 페이지를 찾아야 한다는 단점을 가진다.

이와 같은 단점을 보완하기 위해서 본 논문에서 제안하는 해결 방안은 다음과 같다. 우선 데이터 연동을 최대한 활용하기 위해서 HTML을 XML로 변환해서 데이터와 표현부를 분리한다. 즉, 데이터는 XML로 생성하지만 표현부는 XSL을 이용하면 하나의 원본 데이터를 다양한 디바이스에 적합한 형태로 서비스가 가능하다. 물론 기존 웹 콘텐츠가 XML로 구성되어 있다면 이런 단계가 필요 없지만 현재 대부분의 웹 페이지는 HTML로 구성된 경우가 대부분이며 웰폼드(Well Formed)형식이 아니기 때문에 콘텐츠의 재사용이 용이한 XML의 장점을 살리지 못한다. 둘째, 각기 다른 무선단말기의 특성을 최대한 지원하기 위해 무선 서비스 요청 시그 단말기의 특성을 파악하고 그 단말기에 최적화 구성된 페이지를 서비스한다. 셋째, 일반 PC에서 사용자 관심 영역을 고려한 무선페이지를 서비스 받기 위해 RSS 형식의 구조를 이용한다. RSS를 이용하면 반드시 무선단말기로 접속하지 않아도 자신이 원하는 정보를 채널별로 쉽게 구할 수 있다는 장점을 갖는다. 즉, 하나의 원본을 다양한 방식으로 표현하기 위한 방법을 위해 RSS기반 무선 페이지를 자동 생성하며 이때 자동 생성되는 형식은 RSS의 규약에 따른 XML 페이지이므로 유무선 환경에서도 쉽게 재사용이 가능하도록 한다. 넷째, 사용자의 관심 영역이 될 부분을 다양한 채널로 구성하고 사용자가 자신이 원하는 채널을 선택하도록 해서 불필요한 정보는 배제하고 원하는 정보를 쉽고 빠르게 검색 할 수 있도록 한다.

본 논문의 구성은 다음과 같다. 2장에서는 HTML페이지에서 특정 부분을 RSS XML 데이터로 변환하는 과정과 내용에 대한 설명을 하고 3장에서는 변환된 페이지의 적합성과 서비스의 성능 평가와 내용에 대해 설명하고 실험 결과에 대한 기술을 한다. 마지막으로, 4장에서는 본 논문의 결론과 향후 계획에 대하여 기술한다.

2. 시스템 구성

RSS는 웹사이트를 통한 정보 교환 커뮤니티의 표준으로 자리 잡고 있으며 성공적인 XML서비스로 대표되고 있다. 정보가 공개된 사이트라면 사용자는 이메일이나 사이트를 방문하지 않고도 자신이 원하는 정보를 쉽게 습득할 수 있기 때문에 여러 사이트에서 RSS를 제공중이며 그 확장범위는 더 넓어지고 있다. 대부분의 HTML문서는 XML규약을 지키지 않아 웹폼드하지 않은 경우가 대부분이기 때문에 XML문서로의 변환 과정이 필요하다. 본 연구에서는 모든 HTML에 대한 XML변환은 고려하지 않는다 왜냐하면 무선환경 서비스를 위해서 모든 페이지를 서비스하기 보다는 사용자의 관심 영역만 서비스하는 것이 전송속도나 검색 효율에 있어서

훨씬 적합하기 때문이다. 예를 들어 1024×768의 해상도에 최적화된 웹 페이지를 128×160 크기의 무선 단말기에 서비스 하기 위해서는 신술 비교만으로도 무선페이지가 40페이지 가량 생성 되어야 한다. 따라서 무선 단말기의 특성을 고려한다면 어떤 웹 페이지중 관심 대상이 되는 영역만을 변환해 서비스를 하는 것이 이동단말기의 특성상 더 효율적이다.[10]

제안 시스템의 구성은 (그림 1) 과 같다. 웹 페이지는 XML 구조로 노트 간의 상관관계에 의해 관심영역(Concern Area, CA)을 설정 하게 된다. 이때 설정된 관심 영역은 하나의 채널로 구성된다. 구성된 채널은 무선 단말기의 서비스 요청시 서비스 되는데, 무선 단말기 접속시 웹서버의 호출관리자는 접속 단말기의 특징 정보를 추출해서 XSL을 생성해서 최적화 서비스를 한다. 사용자가 관심 부분을 선택하게 되면 사용자는 페이지의 상세 정보가 담겨있는 연결페이지(Linked Page, LP)에서 광고나 불필요한 부분을 제외한 실제 정보부분만 XML데이터로 구성한다. XML과 XSL은 XSLT를 이용해서 이동 단말기에 서비스되어 정보 요청자는 원하는 내용을 서비스 받을 수 있다. 또한 일반 PC로 접근하는 사용자는 RSS 규약에 맞게 구성된 XML페이지를 이용해서 RSS 채널에 등록해서 최신 정보를 이용할 수 있다.

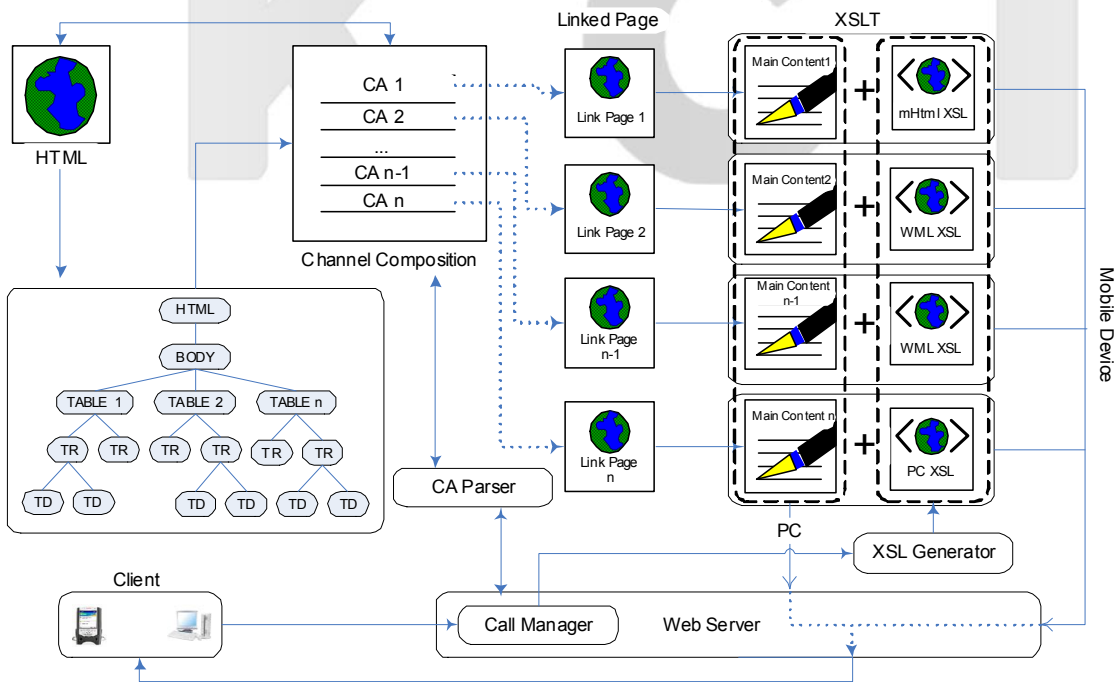


그림. 1 시스템 구조
Fig. 1 System structure

알고리즘 1. 관리자의 채널 설정

- (1) 정보 제공을 위한 웹 페이지 s를 선택한다.
 - (2) k번째 테이블의 x번째 행의 y번째 열의 셀을 선택해서 CA를 설정한다.
- $$f(ca) = S_u + T_k + C_x + R_y$$
- (S_u : Site Url, T_k : Table k, C_x : Column x, R_y : Row y) (3) (3)
- 단계를 반복하며 원하는 CA를 n개 추가해서 기본 채널을 설정
 - (4) CA는 XML구조의 템플릿으로 저장되며 사용자 요청 시 처리되는 기본 파일이다.

제안 시스템의 경우 관리자의 채널 설정 단계와 사용자가 서비스 이용하는 단계로 나누어지기 때문에 다음의 알고리즘1과 알고리즘2 단계를 거치게 된다.

알고리즘 2. 사용자의 정보이용

- (1) 무선 단말기나 PC접속 시 호출관리자는 사용자 단말기 특성 정보를 파악한다.
- (2) 단말기의 정보를 이용해서 XSL 생성기에서는 요청단말기에 최적화된 XSL 파일을 생성한다.
- (3) CA 분석기는CA 템플릿을 분석해서 정보를 요청하는 채널의 링크 페이지 중 실제 중요 데이터가 있는 부분을 XML 데이터로 변환, 이때 XML파일은 RSS 규약을 따른다.
- (4) XSLT를 이용 XML은 XSL과 결합되어 무선 단말기로 콘텐츠를 서비스한다
- (6) PC접속의 경우 (3)단계의 RSS를 사용해서 RSS서비스를 받을 수 있다.

2.1 표현과 데이터 분리

일반 웹페이지는 HTML로 구성되어 있으며 HTML의 특성상 표현부와 데이터를 구분하지 않고 사용된다. 만일 표현과 데이터가 분리된 형태로 존재한다면 하나의 데이터를 다양한 표현을 이용해 다양한 단말기에 서비스가 가능할 것이다. 하지만 기존 웹페이지는 표현과 데이터를 분리해 놓지 않기 때문에 표현과 데이터가 분리된 장점을 살리기 위해서는 전체 페이지를 새롭게 만들어야하는 부담이 생기며 HTML은 나름대로의 장점이 있기 때문에 대부분의 사이트는 HTML을 사용한다. 기존 사이트를 새로 만들지 않고 필요 부분에 대해서만 표현과 데이터를 분리 시키면 표현만 변경함에 따라 많은 수의 무선 단말기에 최적 서비스를 할 수 있다는 장점을 갖는다. 따라서 데이터 저장을 위한 XML과 표현을 위한 XSL(Extensible Style Language)을 생성한 후 XSLT를 이용해 서비스하면 접속 단말기를 위한 실시간 최적 변환 서비스를 수행 할 수 있다.[11]

2.2 기능 및 구조

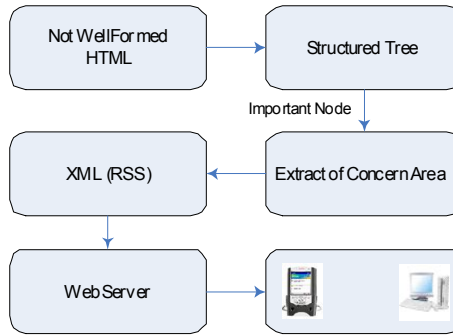


그림 2 기본 흐름
Fig. 2 Basic flow

(그림 2)는 제안 방법의 기본 흐름을 나타낸다. 일반 웹 페이지의 경우 웰폼드 형식이 아니기 때문에 웰폼한 XML 파일로 생성을 위해 현재 웹 페이지의 전체 구조를 파싱해서 트리형태로 구성한다. 이때 비구조적 HTML을 완벽하게 구조적으로 생성하는 것은 문제가 있기 때문에 중요노드에 대한 정보를 추출한다. 본 논문에서 중요노드라고 판단하는 기준은 다음페이지로의 링크의 유무를 이용해 판단한다. 대부분의 사이트의 경우 섹션별로 콘텐츠를 구성하고 있으며 이를 메인 페이지라고 할 경우메인 페이지는 요약해 제공하고 연결된 하위 페이지는 메인 페이지의 상세 정보를 담고 있다. 즉, 하위 페이지로의 링크를 갖는 텍스트나 이미지의 경우 하위 페이지에 상세한 정보를 더 포함하고 있기 때문에 메인 페이지에서 다른 부분에 비해서 중요한 내용을 포함한다고 볼 수 있다. 따라서 (그림 3)에서와 같이 현재 페이지에서의 링크가 걸린 텍스트와 이미지 요소를 추출한다. 오른쪽 창은 현재 페이지의 노드 트리를 보여주고 있다.



그림 3 링크 노드 추출
Fig. 3 Extract a linked node

페이지 관리자는 현재의 페이지에서 무선페이지로 서비스 할 관심 영역을 설정해서 무선페이지로 서비스할 대상 영역 설정을 한다. 대개의 웹 페이지는 구조나 관계의 유지를 위해 <table>, <div>, <iframe> 등의 태그를 이용해 페이지의 구조를 디자인하며 일반적으로 한 사이트의 구조는 전체적으로 바뀌지 않는 이상 거의 같은 구조를 유지한다. 따라서 실제 정보를 포함하는 태그를 이용해 전체 페이지의 구조를 파악할 수 있다. 몇 번째 테이블의 몇 번째 행, 혹은 하나의 중요노드를 선택했을 때 그 노드를 포함하는 전체 테이블을 파악한다. 이렇게 관리자에 의해 선택된 페이지는 하나의 채널로 구성되며 관리자는 다수개의 채널을 생성할 수 있다. 편집된 각 채널은 채널 자료구조로 저장한다. 이 자료구조에는 관리자가 선택한 정보의 위치정보를 담고 있으며 이 자료구조를 이용해 XML형식의 RSS파일을 생성한다. 이렇게 생성된 RSS파일은 무선단말기가 서비스를 요청할 때나 일반 PC의 서비스 요청시 사용된다. 무선단말기에 서비스 될 때 RSS파일은 XML 데이터의 역할을 해서 무선단말기에 최적화된 XSL과 결합되어 각 단말기에 맞는 마크업 언어를 생성 서비스한다. 일반 PC로 서비스 될 경우는 푸시(Push)서비스가 가능한 RSS파일의 역할을 수행한다. RSS를 이용할 경우 정확하고 빠르고 사용자의 정보선택이 가능한 형태로 서비스 받을 수 있다는 장점을 가지게 된다.

2.3 채널 구성 템플릿

웹 페이지에서 원하는 부분의 정보를 제공하기 위한 채널 정보는 XML형식의 템플릿(Template) 파일로 구성된다. 관리자가 채널 구성을 원하는 중요노드를 선택해서 구성할 때 템플릿 내부에 저장되는 내용은 HTML문서 상에서 몇 번째 테이블의 몇 번째 컬럼에 위치한 노드이며 어떤 문서와 연결되었는가에 대한 링크 정보 등의 필요한 정보를 저장한다.

이 템플릿 파일은 한번만 생성해 두고 계속 사용되며 단말기 접속 시 템플릿 정보에 의해 HTML파일의 내용정보를 추출하는데 사용된다. 따라서 이 템플릿 파일에 의해 사용자는 최신의 정보를 실시간 전달 받을 수 있다.

채널구성된 템플릿과 템플릿을 통한 변환 식은 (1)과 (2)와 같이 표현된다. 어떤 웹 콘텐츠 S_i 를 채널로 구성하기 위한 요구되는 시간 t 의 변환 시간 함수를 Ω 로 정의하고 Ψ 는 변환된 콘텐츠 S_i 의 실제 서비스 시간 t 의 함수로 정의한다.

$$\Omega(S_i, t) = T_{(cg)} + T_{(sr)} + T_{(bw)}$$

(t : ServiceTime, cg : ChannelGenerate, wr : Serverresponse, bw : Bandwidth) (1)

$$\Psi(S_i, t) = T_{(fms)} + T_{(pg)} + T_{(xg)} + T_{(RSS + XSL)} + T_{(bw)}$$

(t : ServiceTime, fms : FindMobileSpec, pg : PageGenerate, xg : XSLGenerate, bw : Bandwidth) (2)

따라서 채널 구성을 위한 페이지 S_i 의 생성시간은 페이지를 분석하고 파싱해서 템플릿 채널을 구성하는 시간과 웹서버의 응답시간 그리고 데이터 전송 대역폭의 합으로 나타낼 수 있으며 이때 생성에 소요되는 시간 t 는 작을수록 유리하다. 변환된 페이지 서비스를 위한 실제 서비스 시간은 무선단말기 접속시 특성 파악과 XSL 생성, 템플릿 채널에서 페이지 생성하는 시간과 XSLT로 결합되어 무선단말기로 서비스하는 시간이 변환된 콘텐츠의 실제 서비스 시간 t 로 결정된다.

2.4 호출 관리자(Call Manager)

호출 관리자는 단말기의 종류와 성능, 하드웨어 특성 등을 파악하는 모듈이다. 단말기가 서비스를 요청을 할 경우 접속 단말기의 MIME 헤더 정보를 분석해서 지원 브라우저, 지원 이미지 종류, 디스플레이 크기, 통신 속도 등의 정보를 알아내며 이 정보를 이용해 XSL을 생성하는데 사용된다[3][9].

2.5 XSL 생성기

RSS형식의 XML로 구성된 데이터 부분을 생성한 후 무선 단말기에 최적화된 형태로 서비스하기 위해서는 그 단말기에 최적적으로 표현할 수 있는 XSL이 있어야 하는데 모든 단말기에 최적화된 XSL을 미리 만들어 놓기는 어렵기 때문에 호출관리자에서 찾아낸 접속단말기의 특성에 맞는 최적 XSL을 동적으로 생성 한다[3][9].

2.6 선택적 최적화 서비스

HTML 페이지에서 추출된 RSS XML 데이터는 XSL 생성기에서 생성된 XSL과 결합되어 XSLT를 이용해 최종 서비스 가능한 무선페이지로 변환된다. 무선 단말기가 지원하는 브라우저와 디스플레이 크기 등의 특성을 고려해서 XSL이 생성되었기 때문에 생성되는 페이지는 접속 단말기에 최적화된 형태로 서비스 된다. 만일 접속 단말기가 일반 PC라면 정보의 푸시 서비스가 가능한 RSS의 특성을 그대로 이용할 수 있다.

3 변환된 페이지의 적합성과 서비스시의 성능 평가와 내용

3.1 실험 환경

실험 환경은 다음과 같다. 운영체제는 Windows XP Professional, 웹서버 IIS 5.0, CPU는 펜티엄4 3.0Ghz, 메모리 1G, 개발도구는 Visual Studio 6.0을 사용했다. 클라이언트 접속을 위해 OpenWaveSDK와 휴대폰, PDA를 이용한 테스트를 수행했으며 SharpReader RSS 리더기를 이용해 PC에서 RSS 리더를 테스트했다.

3.2 웹 콘텐츠 정보 분석

<표 1>에서는 국내와 해외의 유명 사이트의 메인페이지에 대한 분석표이다. 국내 사이트의 경우 해외사이트에 비해 많은 링크정보와 이미지를 사용하기 때문에 페이지의 크기가 크며 이정도 크기의 페이지를 서비스하기 위해서는 무선단말기의 처리용량이나 데이터 전송속도가 빨라져야 한다. 또한 무선 단말기의 작은 디스플레이는 이 모든 데이터를 표시하고 검색하기에는 많은 무리가 따른다.

표 1 웹 콘텐츠 정보

Table. 1 The information of web contents

	텍스트 링크	이미지 링크	전체 이미지	테이블 개수	크기 (KB)
www.naver.com	375	43	59	23	210
www.empas.com	386	83	101	42	286
www.yahoo.com	171	13	32	9	109
www.nbc.com	150	47	56	58	360

3.3 실시간 콘텐츠 생성 및 접근

<표 2>는 각 사이트 당 2개의 뉴스 서비스 채널에 대한 생성 시간과 서비스 시간 그리고 정확도에 대한 분석 표이다. 생성 시간은 관리자가 서비스 할 뉴스 섹션과 테이블 선택을 한 후부터 계산 했으며 접근 시간은 무선단말기의 서비스 요청 시작부터 사용자가 원하는 내용을 선택하고 결과 화면이 표시되는 시간까지를 체크했다. 사용자가 원하는 내용을 선택하는데 걸리는 시간을 최소화하기 위해 가장 첫 번째 기사를 선택했지만 어느 정도의 오차는 있다. 위와 같은 테스트를 각 10번씩 반복 수행했으며 원본 사이트의 뉴스기사가 업데이트 된 후, 무선단말기에도 실시간 업데이트가 수행되어야 하기 때문에 실시간 반영율을 각 10번씩 테

스트한 결과를 <표 2>에 나타내고 있다. 그리고 웹 브라우징 후 인터넷 캐쉬파일은 제거하지 않고 그대로 실험에 수행했다. PC의 RSS리더기를 이용한 RSS 서비스 테스트의 경우 정확도와 실시간 반영율은 <표 2>와 동일한 결과를 나타낸다. 이는 생성된 RSS 파일을 유무선단말기에 동일하게 사용하기 때문이다.

표 2 테스트 결과
Table. 2 The result of the test

	생성 시간 (ms)	접근 시간 (ms)	정확도 (%)	실시간반영율 (%)
www.naver.com	488	512	98	100
www.empas.com	512	613	97	100
www.yahoo.com	613	728	97	98
www.nbc.com	884	1015	95	97

접근시간은 변경된 기사의 반영 위해서 무선단말기의 서비스 요청시마다 하위 기사 페이지를 새로 생성하기 때문에 템플릿 페이지 생성 시간과 비슷하게 소요된다. 이러한 부분은 뉴스 페이지의 변경이 생기지 않았을 경우에는 새롭게 페이지를 생성하지 않고 기존의 변환된 내용을 서비스하면 더 시간을 줄일 수 있을 것이다. 좀 더 효율적인 페이지 관리와 서비스를 위한 알고리즘과 정확도를 더 높이기 위한 개인화 기법과 데이터마이닝 기법에 대한 연구는 진행 중이다.

(그림 4)와 (그림 5)는 테스트 결과 화면을 나타내고 있다. (그림 4)에서 사용자가 원하는 기사를 선택하면 (그림 5)와 같은 실제 기사를 표시한다. (그림 4)의 원본 HTML의 크기는 이미지등의 멀티미디어 데이터가 포함된 209KB였지만 실제 뉴스 부분만 변환된 경우 텍스트 정보만 포함하고 있기 때문에 493B로 약 1/400 크기이며 (그림 5)의 원본 콘텐츠의 크기는 317KB인데 무선단말기에 서비스되는 크기는 2.6KB이므로 약 1/100의 크기이다. 테스트 결과 뉴스기사 선택 페이지의 경우는 평균 1/400, 실제 뉴스 정보의 경우 평균 1/130 정도의 크기가 되었다. 따라서 PC에 비해 협소한 대역폭과 성능을 가진 무선단말기의 특성에 적합하게 서비스 된다고 볼 수 있다.



그림. 4 뉴스 기사 선택
Fig. 4 News index page



그림. 5 실제 기사 내용
Fig. 5 Real news contents

4. 결론

본 연구는 기존의 웹 콘텐츠의 변경 시 변경된 내용을 수작업으로 재 가공 없이 무선단말기와 RSS를 이용해 실시간 서비스가 가능하다는 점에서 의미가 있다. 하지만 변경된 웹 콘텐츠를 무선페이지로 실시간 변환하기 때문에 웹 서버에 부하가 걸릴 수 있다는 점은 문제로 남을 수 있다. 하지만 실시간 감시 프로세스가 웹 콘텐츠 업데이트시 미리 페이지 정보를 분석하고 무선 페이지를 생성해 놓고 매번 새로운 페이지를 생성하는 것이 아니라 생성되었던 페이지를 재사용한다면 이런 문제는 해결될 수 있을 것이다. 그리고 사이트의 중요 구조가 바뀌면 서비스가 불가능하다는 단점을 가지고 있지만 구조가 바뀌는 일은 흔하지 않고 바뀔 때 마다 관리자가 한번씩 채널을 새롭게 설정해 주면 자동적으로 무선단말기에 서비스가 가능하기 때문에 큰 문제가 되지는 않는다. 이와 같은 방법을 이용해 원하는 콘텐츠만 서비스 될 경우 크기는 평균 1/150정도로 줄일 수 있으며 서비스 시간 또한 같은 네트워크 속도를 갖는다면 같은 비율로 빠른 서비스가 가능하다. 또한 콘텐츠 생성자는 수동 작업을 수행하는 경우에 비해 작업시간을 약 40%정도 줄일 수 있으며 실제 서비스가 될 때는 자동 처리가 되므로 서비스 제공자의 소요시간은 거의 없다.

본 연구를 적용하면 이미지나 동영상등 멀티미디어에 데이터에 대한 변환과 서비스도 가능할 것으로 생각되며 차후 과제로는 현재 진행중인 멀티미디어 적용 변환연구를 적용해서 멀티미디어 데이터들도 쉽게 사용할 수 있도록 할 예정이다.

참고문헌

- [1] [1] <http://www.openmobilealliance.org/>
- [2] <http://www.netscape.com/>
- [3] <http://www.w3.org/RDF/>
- [4] 류동엽, 한승현, 이근수, "멀티플랫폼 이동단말기의 무선 웹 서비스를 위한 시스템 설계 및 구현", 한국 컴퓨터 정보학회 논문지 10권 4호, pp. 355-364, 2005
- [5] Yonghyun Hwang, Jihong Kim, Eunkyong Seo, "Structure-Aware Web Transcoding for Mobile Devices", IEEE Internet Computing, vol. 7, no. 5, pp.14-21, Sep.Oct. 2003.
- [6] S. Lok and M. Kan. Employing Natural Language Summarization and Automated Layout for Effective Presentation and Navigation of Information Retrieval Results. The 12th International World Wide Web Conference, Budapest, Hungary, 20-24 May 2003.
- [7] T. Bickmore, A. Girgensohn, and J.W. Sullivan, "Web Page Filtering and Reauthoring for Mobile Users," Computer J., vol. 42, no. 6, pp. 534-546, 1999
- [8] T. Bickmore and W. Schilit, "Digester: Device-Independent Access to the World Wide Web," Computer Networks and ISDN Systems, vol. 29, no. 8, pp. 1075-1082, 1997.
- [9] 류동엽, 한승현, 임영환, "편리한 무선 인터넷 콘텐츠 생성을 위한 TransGate 시스템", 한국인터넷 정보학회 논문지, 7권 4호, pp.37-52, 2006.
- [10] Yu Chen, Xing Xie, Wei-Ying Ma, Hong-Jiang Zhang, "Adapting Web Pages for Small-Screen Devices", IEEE INTERNET COMPUTING, January February, pp.2-8, 2005.
- [11] Timo Laakko, Tapio Hiltunen, "Adapting Web Content to Mobile User Agents", IEEE INTERNET COMPUTING, march april, pp.46-53, 2005.

저 자 소개



류 동 업

2003년 : 송실대학교
컴퓨터공학과 졸업(공학석사)
2003 ~ 현재 : 송실대학교 컴퓨터학과
박사과정



한 승 현

2002년 8월: 송실대학교
컴퓨터공학과 졸업 (공학석사)
2003년~현재: 송실대학교 컴퓨터학과
박사과정



임 영 환

1977년 : 경북대학교
수학과 졸업(이학사)
1979년 : 한국과학 기술원 전산학과
(이학석사)
1979년~1996년 : 한국 전자통신연구소
책임연구원
1985년 Northwestern University
(이학박사)
1996년~현재 : 송실대학교 미디어학부
교수

