

사전기반 웹메일 문서 생성기

최지연, 민수홍, 조동섭
이화여자대학교 과학기술대학원 컴퓨터학과

Dictionary based Web Mail Document Generator

Ji-Yeon Choi, Su-Hong Min, Dong-Sub Cho
Dept. of Computer Engineering & Science, Ewha Womans University

Abstract - WWW(World Wide Web)은 인터넷상에 광범위하게 분산된 다양한 형식의 데이터를 손쉽게 검색할 수 있도록 하며, 모든 응용의 사용자 환경을 손쉽게 포함시킴으로써 새로운 미디어 타입의 통합을 위한 가장 성공적인 기술로 평가받고 있다. 처음 Tim Berners-Lee에 의해 처음 개발될 당시 WWW는 텍스트 기반의 구조에 맞추어 비동기적인 형태의 검색과 상호작용을 지원하도록 설계되었다. 반면 최근 인터넷상의 상당수를 차지하는 오디오, 비디오와 같은 실시간 환경의 멀티미디어 데이터들은 연속적, 동기적이며, 대화형 특징을 포함하고 있어 WWW에서 이러한 미디어들을 통합하기 위해서는 구조적인 확장이 요구된다. 본 논문에서는 현재 가장 많이 쓰이는 WWW 기능인 e-Mail 서비스에 대한 동기적 방법론을 제안하고자 한다. e-Mail 사용자가 HTML tag에 대한 사전 지식이 없이도 사전을 기반으로 e-Mail의 형식을 지정할 수 있도록 하며 이렇게 함으로써 보다 동기적인 형태의 서비스를 구현하도록 한다.

1. 서론

텍스트 기반의 하이퍼텍스트 개념에서 출발한 WWW 기술은 새로운 미디어 타입을 통합하기 위한 가장 성공적인 기술로 평가받고 있다. Tim Berners-Lee에 의해 처음 개발될 당시 WWW의 텍스트 기반의 구조에 맞추어 HTTP(HyperText Transfer Protocol), HTML(HyperText Markup Language), URL(Uniform Resource Locator)이 제안되어 이를 통해 서버와 클라이언트간의 비동기적인 형태의 검색과 상호작용을 지원하도록 설계되었다. HTML은 1990년 이후 웹 상에서 가장 많이 사용되는 문서 교환의 표준 형식으로서 1986년에 제정된(ISO 8879) SGML(Standard Generalized Markup Language)를 바탕으로 정의된 하나의 응용이다. HTML은 구조적 특성을 갖는 이유로 손쉬운 편집이 어려웠으나 웹이 전 세계 사용자들의 각광을 받으면서 HTML을 이용한 문서편집이 활발해지고 빠르고 간편한 WYSIWYG(What You See Is What You Get) 방식의 HTML 편집기도 나오게 되었다[1].

그러나 HTML은 단순한 텍스트 기반의 구조를 통한 비동기적인 형태의 서비스 유형으로 문서검색과 단순한 표현 방식을 제공한다. 따라서 다양한 형태의 동기적인 데이터 및 네트워크 자원을 효과적으로 교환 및 검색하기에는 한계를 가지고 있다. 특히 WWW의 대표적인 기능이라고 할 수 있는 e-Mail 서비스는 대부분의 인터넷 사용자들이 사용하는 기본적인 기능임에도 불구하고 비동기적 서비스를 유지하고 있으며, 모든 사용자들이 HTML tag를 알고있다는 전제하에 제한된 동기적 서비스를 제공하고 있다. 그러므로 e-Mail 사용자가 보내고자 하는 메일의 본문을 편집할 때, 사용자가 HTML의 tag에 대한 지식이 전혀 없을 경우, 사용자는 단순한 텍스트 형식의 메일만을 전송할 수 있다.

그러므로, 본 논문에서는 효과적인 e-Mail 사용을 위해, 사용자가 HTML tag에 대한 지식이 전혀 없이도 표현하고자 하는 개체에 대한 정보가 사전에 미리 등록되어 있다면 e-Mail의 내용을 작성할 때, 표현하고자 하는 개체 정보가 자동으로 삽입되어 효율적으로 문서가 생성되도록 한다.

본 논문은 2장에서 관련 연구에 대해서 기술하며, 3장에서는 제안하는 메일 서비스에 대해 설명한다. 4장에서는 본 논문의 결론을 맺고, 향후 연구 계획에 대해 언급한다.

2. 관련연구

2.1 메일의 전송과정

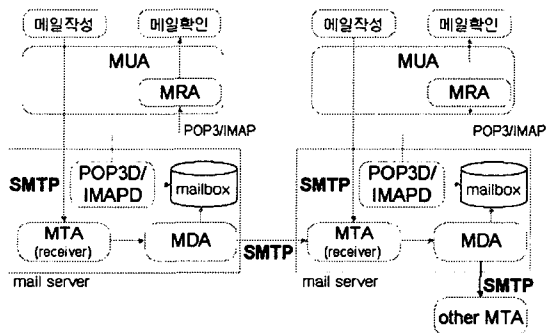


그림1. 메일의 전송 과정

메일을 전송하기 위해서는 사용자가 일반적으로 MUA (Mail User Agent)에 포함되어 있는 메일 작성 프로그램을 사용하거나, 직접 메일 서버의 MUA에서 메일을 작성해야 한다[2]. MUA는 사용자가 메일을 송수신할 때 사용하는 클라이언트 프로그램으로, 마이크로소프트사의 아웃룩 익스프레스나 리눅스, 유닉스에서 사용하는 mail 프로그램을 예로 들 수 있다. MUA는 사용자가 작성한 메일을 메일서버의 MTA(Mail Transfer Agent)로 전송한다. MTA는 인터넷상에 있는 하나의 컴퓨터로부터 다른 컴퓨터(수신자측의 메일 서버)로 메일을 전달하는 프로그램이다. MTA는 수신된 메일의 수신자가 로컬 메일 서버의 사용자인지, 아니면 다른 메일 서버의 사용자인지를 판단한다. 만약 로컬 사용자인 경우에는 해당 MDA(Mail Delivery Agent)를 통해서 로컬 파일 시스템에 있는 사용자의 우편함에 메일을 저장한다. MDA는 메시지를 사용자의 우편함에 쓰기위해 MTA가 사용하는 프로그램이다. 그리고 로컬 사용자가 아닌 경우에는 별도의 MDA를 사용해서 다른 MTA로 메일을 전달한다. 메일을 수신받은 수신자의 MTA는 다시 로컬 메일서버의 사용자인지를 판단한 후에, MDA를 통해서 파일 시스템의 사용자 우편함에 메일을 저장한다. MTA와 MDA는 sendmail[3]이나 qmail[4]처럼 일반적으로 하나의 메일 서버 어플리케이션에 통합되어 있다. 사용자

이 논문은 2003년도 두뇌한국21사업에 의하여 지원되었음

MUA에서 메일 서버로 메일을 전송할 때와, 메일의 수신자가 로컬 사용자가 아니라서 다른 메일 서버에 있는 MTA로 메일을 전송할 때 SMTP[5]를 사용한다.

사용자가 자신의 메일을 확인하기 위해서는 MUA에 포함되어 있는 MRA(Mail Retrieval Agent)를 사용해야 한다. MRA는 원격지 서버에 있는 우편함으로부터 사용자의 MUA로 메시지를 가져오는 서비스를 제공하는 프로그램이다. MRA가 메일 서버에 있는 POP3 데몬에 접속하면, POP3 데몬은 사용자의 우편함에서 메일을 읽어, MRA에게 전달한다. 사용자의 우편함으로부터 메일을 읽어올 때는 POP3를 주로 사용하지만, 요즘에는 많은 네트워크 지향적인 옵션을 제공하는 IMAP(Internet Message Access Protocol)의 사용이 선호되고 있다.

2.2 전자우편 시스템

전자우편 클라이언트 프로그램으로 SMTP 서버에 편지를 보낸다. 전자우편 메시지 수신인의 도메인이 서버가 있는 도메인과 같으면 아스키 텍스트를 서버에 저장하고, 자기 도메인이 아닌 경우 해당 서버에 전달한다. 전자우편 메시지가 수신자의 SMTP 서버에 도착하면 수신자에 할당된 사서함에 저장된다. 수신자는 아웃룩 익스프레스 같은 우편 클라이언트에서 POP3 프로토콜로 서버에 접속해서 자신에게 온 편지를 가져간다. 이 과정에서 포트25를 사용하는 SMTP로 편지를 서버에게 전달하고, 포트110을 사용하는 POP3 프로토콜로 편지를 클라이언트까지 배달한다. 그림2는 전자우편 배달과정을 보여준다[6].

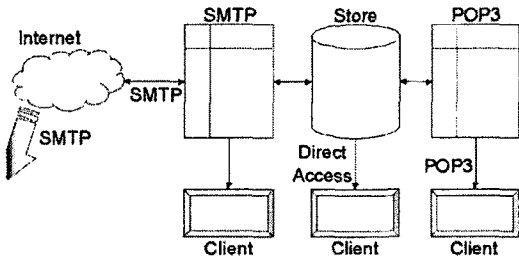


그림2. 전자우편의 전송

2.3 e-Mail 전송 프로토콜

SMTP	<ul style="list-style-type: none"> TCP/IP 계열의 표준화된 통신규약. 메일 서버들 사이의 전송. 메일 서버와 메일 클라이언트 사이의 전송(UNIX). 메일을 사용자의 시스템에 저장하지 않는다. 메일 저장소에 따로 저장. 사용자의 요구시 가져온다.
POP3	<ul style="list-style-type: none"> 메일 서버와 메일 클라이언트 사이의 전송(PC). 전송될 메일을 사용자의 시스템에 저장. 사용자의 메일 조작이 용이.
IMAP	<ul style="list-style-type: none"> 서버에 기반을 둔 메일 저장. 메일의 머리말만을 복사. POP3의 기능을 확장하여 대체할 목적으로 개발.

표1. e-Mail 전송 프로토콜

3. 사전기반 웹 메일 문서 생성기

기존의 메일 서비스는 HTML의 비동기적인 형태의 서비스이며 사용자가 HTML tag에 대한 지식을 갖고 있다는 전제하에서 동적인 서비스를 제공하고 있다. 그러므로 HTML tag에 대한 지식이 전혀 없는 사용자들은, 인터넷의 상당부분을 차지하는 오디오, 비디오와 같

은 연속적이고 동기적인 멀티미디어 데이터들을 메일 본문에 삽입하고자 할 때, 큰 불편함을 느끼게 된다.

따라서 HTML tag에 대한 지식이 전혀 없는 사용자들도 웹 문서에 원하는 정보를 삽입시킬 수 있도록하며, 일반 사용자들도 보다 편리하게 웹 문서를 다룰 수 있도록 하는 동기적인 형태의 메일 서비스가 필요하다.

웹 메일 문서를 생성하기 위해서는, 먼저 개체들을 각각의 성격이나 형식 등을 기준으로 분류하여야 하며, 이때 분류된 각 개체들은 사용자에 의해 사전으로 저장된다. 사용자가 메일을 작성하여 보내면, 사용자가 작성한 문서는 전송되기 이전에 사용자가 설정해놓은 사전을 거친다. 사전을 통해 HTML 문서는 재생산되어 목적지로 전송된다.

3.1 TXT2HTML 전처리기의 구성

전체적인 구성을 살펴보면, 사용자가 사전에 등록된 정보는 Dictionary DB에 저장된다. 사용자가 메일을 작성하면 메일은 전송되기 전에 DB에 저장된 정보를 바탕으로 재생산되어 전송되며 각 단계에 따른 처리과정은 다음과 같다.

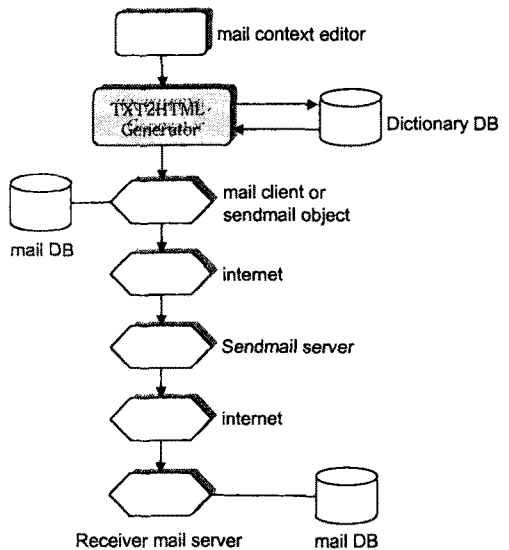


그림3. TXT2HTML Generator

3.2 전체 구조

사용자는 메일을 일반 텍스트 형식으로 작성한다. 사용자가 메일을 전송할 것인지를 결정하면, 메일은 사용자가 미리 만들어 둔 Dictionary DB를 거치게 된다. 텍스트 문서는 이 Dictionary DB안의 정보를 바탕으로 텍스트 문서의 내용과 Dictionary DB 안의 목록이 일치하는 것이 있는지 찾는다. 일치하는 정보가 검색되면 DB에 저장된 목록들의 정보들을 사용자가 작성한 텍스트 메일에 첨부하여 새로운 HTML 문서를 생성한다. 이에 대한 상세 과정은 그림4와 같다.

4. 결 론

본 논문에서 제안하는 사전기반 웹 문서 생성기는 텍스트 형식의 문서뿐만 아니라, 모든 형식의 문서를 읽어서 사용자가 생성한 Dictionary DB에 저장된 정보와 일치하는 정보가 있을 때 기존의 문서에 새로운 내용을 추가하여 HTML 형식의 문서로 바꿔주는 기능을 한다. 이를 이용하여, 사용자는 HTML tag에 관한 지식이 전혀

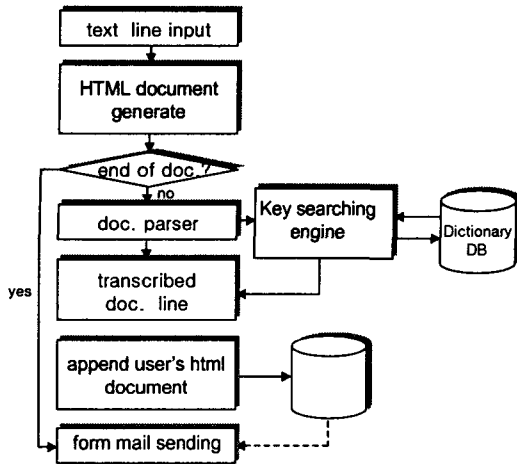


그림4. e-Mail 문서 생성과정

없이도 보내고자 하는 메일 내용에 색 기능과 하이퍼링크 기능 등을 추가시킬 수 있으며 더 나아가 미디어나 오디오와 같은 동적인 형태의 멀티미디어 기능들도 삽입시킬 수 있다. 기존 문서상의 키워드에 관련된 다양한 정보들을 자동으로 연결시켜 웹 문서 작성이 편리하도록 하며, 새로운 통합 정보를 구축하는데 용이하도록 한다.

향후 연구과제는 현재 텍스트 단계에까지만 그친 웹 문서 생성기의 기능을 소리, 동영상 및 가상현실과 같은 정보들과도 연결시켜, 멀티미디어 개체들을 표현할 수 있도록 하이퍼 미디어 기능을 지원하는 메일 시스템을 구축하는 데에 있다.

[참 고 문 헌]

- [1] w3 team at CERN, The WWW Book, 1995
- [2] David Wood, "Programming Internet Email", O'reilly, pp.18-20, June 1999
- [3] sendmail 18.11.0, <http://www.sendmail.org>
- [4] Dan Bernstein, qmail-1.03, <http://qmail.org>
- [5] Jonathan B.Postel, "Simple Mail Transfer Protocol" in RFC821, August 1982
- [6] Leon Braginski and Matt Powell, "Running Microsoft Internet Information Server," Microsoft Press, 1998.