

메타 전자 쇼핑몰의 설계 및 구현

신민화, 최한석
목포대학교 정보공학부 멀티미디어전공
e-mail:minandjih@hanmail.net

Design and Implementation of Meta Electronic Shopping Mall

Min-Hwa Shin, Han-Suk Choi
Dept of Computer Science, MokPo National University

요약

최근 전자상거래의 상업성이 부각되면서 많은 기업들과 개인사업자들이 독자적인 웹서버를 구축하고 자사의 상품 또는 기업에 대한 광고 내용을 인터넷을 통해 서비스를 하고 있다. 그러나 소비자에게 홍보가 부족하고, 상품의 종류와 정보의 부족으로 외면을 당하고 있는 웹사이트들이 많이 있다. 본 논문에서는 이러한 기업들의 카탈로그를 수집하여 재구성한 후, 한 사이트에서 여러 기업의 상품을 검색할 수 있도록 하여 소비자의 접근과, 각 기업의 상거래 사이트를 활성화 할 수 있는 메타 전자 쇼핑몰 개발에 관해 기술한다. 본 시스템의 특징은 로봇이 자동으로 데이터 베이스를 구축하여 소비자가 쉽게 검색 할 수 있도록 주제어 검색과 디렉토리 검색 인터페이스를 구성을 가지고 있다는 점이다.

1. 서론

최근 인터넷의 급속한 확대와 함께 화제를 모으고 있는 것이 전자상거래(EC; Electronic Commerce)이다. 특히 월드와이드웹(WWW; World Wide Web)의 사용이 일반인에게 확산되기 시작하면서 급속히 발전하였으며, 인터넷 쇼핑몰(Internet Shopping Mall), 전자상점(Electronic Retailing), 가상점포(Virtual Storefront), 온라인 점포(On-Line Shop)등의 다양한 이름으로 불리우고 있다[1].

전자상거래의 치열한 경쟁에서 성공하기 위해서는 소비자들로 하여금 상품에 대한 내용을 담고 있는 문서에 많은 접근을 할 수 있도록 유도하여야 한다. 이런 맥락에서 회사의 이미지 마케팅을 이미 보유하고 있는 대기업들은 유리한 입장에 있다. 그러나 중소기업이나 지방 자치 단체의 토산품 판매 또는 개인 사업자의 입장에서 보면 매우 불리한 편이다.

이러한 문제를 해결하기 위해 전자상거래에서 열악한 입장에 처해있는 기업들의 사이트에서 전자카탈로그 정보를 수집하여 하나의 사이트에서 정보를

제공해주는 메타 전자쇼핑몰의 구축이 매우 중요하다. 메타 전자 쇼핑몰의 사이트에 접근한 고객은 여러 물건을 드나드는 불편함이 없이 하나의 물에서 쇼핑하듯이 편리하게 쇼핑을 즐길 수 있다.

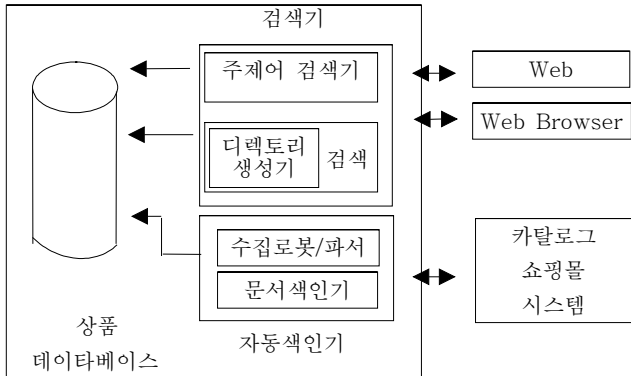
본 논문에서는 전자 상거래 시스템을 구축한 기업들 중에 상품 홍보에 취약한 쇼핑몰들을 하나의 메타 쇼핑몰 사이트로 수집, 가공하여 고객에게 제공함으로써 상거래 시스템의 효율을 높일 수 있는 로봇을 이용한 메타 전자 쇼핑몰 구축 모델을 제안한다. 제 2장에서는 메타전자쇼핑몰 설계 및 구성에 대해 설명하고, 3장에서는 결론을 내린다.

2. 메타 전자 쇼핑몰 설계 및 구성

본 논문에서는 네 개의 지방 자치 단체(목포, 무안, 함평, 강진)의 웹사이트에 등록된 토산품 전자 카탈로그에서 정보를 수집하여 하나의 사이트에서 토산품에 관련된 상품을 고객으로 하여금 구매할 수 있도록 돕는 메타 전자 쇼핑몰을 구현하였다.

로봇이 상품 카탈로그가 있는 사이트에 접근하여

문서를 가져오면 문서 색인기는 주제어를 추출하여, URL과 주제어를 데이터베이스에 등록을 하게 된다. 그러면 사용자는 주제어 검색 또는 디렉토리 인터페이스를 통하여 원하는 정보를 추출하게 된다. [그림 2-1]은 토산넷의 구성도이다.



[그림 2-1] 토산넷 구성도

토산넷 시스템은 크게 자동 색인기와 검색기, 그리고 상품 정보를 저장하는 저장 시스템으로 구성되어 있다. 클라이언트 사용자들은 Netscape나 IE등 기타 범용 웹브라우저를 이용하여 인터넷 환경에서 사용이 가능하다. 서버 시스템은 PⅡ-400, Power Linux 6.0환경에서 구현되었으며 웹 서버로는 Apache Web Server를 이용하였다. Database로는 PostgreSQL 6.5를 이용하였고, DB와 Web Server와의 인터페이스는 PHP3.0을 사용하였다.

2.1 자동 색인기

자동 색인기는 크게 각 자치 단체의 카달로그의 정보를 수집하여, 수집된 문서를 파서하는 수집로봇과 파서기 그리고 파서된 문서의 내용을 인덱스화한 후 데이터베이스에 저장하는 문서 색인기로 나눌 수가 있다.

수집 로봇은 시스템의 지시에 의해 작동을 시작하여 시작 URL에 접근하여 토산품이 등록된 문서를 가져오게 된다. 이때 수집 로봇은 넓이 우선 방식에 의해 탐색을 시작한다.



[그림 2-2] 무안 한우 카탈로그



[그림 2-3] 함평 춘란 카탈로그

[그림 2-2,3]는 일부 자치 단체의 카달로그 정보이고, [그림 2-4]는 로봇 프로그램의 일부이다.

수집 로봇에 의해 전달된 문서를 받은 파서는 수신 웹 페이지로부터 헤더와 타이틀, 그리고 URL을 추출해 내고, 웹 페이지의 모든 태그를 제거한다. 그리고 이렇게 태그가 제거된 웹 페이지를 문서 인덱서에게 전달한다. 인덱서는 파서로부터 수신한 태그가 제거된 웹 페이지내에서 모든 단어를 분리해 내고, 이 단어를 검색엔진에서 더 빨리 찾을 수 있도록 토산품/특산품과 관련된 단어를 중심으로 가중치를 부여하여 문서에 대한 주제어를 선정하게 된다. 그런 후 문서에서 추출된 주제어, 지방이름 URL을 Database에 등록한다.

해당 웹 서버와 연결을 설정했으면 웹 서버에게 요청 메시지를 보내고, 그리고 나서 응답 메시지를 받아서 상태 코드를 확인하고, 정상적인 웹 페이지라면 해당 웹 페이지를 수신한다.

```
int GetHTMLDoc(char *HostName, GONE URL)
{
    char RequestMessage[REQUEST_MESSAGE];
    char Header[HEADER_LENGTH];
    char url[STATUS_MESSAGE],ch;
    char StatusMessage[STATUS_MESSAGE];
    int Count, Length, TimeOut, Flag=0;, size=0; Status;
    MSG Msg;

    /* 웹서버에게 요청 메시지를 보낸다. 응답 메시지를 HTTP Version
    1.0형식으로 보내 달라고 요청을 한다. */
    sprintf(RequestMessage, "GET %s` TP/1.0WrWn",URL.PathName);
    if(write(SocketFD, RequestMessage, strlen(RequestMessage)) < 0) {
        SendStatusMessage("<Robot> Occurred Error while Sending to
        HTTP Header Wn");
        return(FAIL);
    }

    /* 로봇이 수신할 문서는 텍스트임을 알린다.*/
    sprintf(RequestMessage, "Accept : text/html WrWn");
    if(write(SocketFD, RequestMessage, strlen(RequestMessage)) < 0) {
        SendStatusMessage("<Robot> Occurred Error while Sending to
        HTTP Header Wn");
        return(FAIL);
    }

    /* 웹서버로부터 웹페이지를 수신한다. */
    SendStatusMessage("<Robot> Receiving HTML Document ... Wn");
    TimeOut =0;
    while((Length = read(SocketFD, &ch, 1)) != 0){
        if(ch == TERMINATOR) break;

    /*파서에게 웹 페이지를 전달한다. */
        if(write(HTML_fd, &ch,1) < 0) {
            SendStatusMessage("Wn<Robot> Can` t Send HTML
            Document to ParserWn");
            break;
        }
    }

    /* 웹 페이지를 모두 수신했음을 파서에게 알린다. */
    write(HTML_fd, &TERMINATOR, 1)

    SendStatusMessage("<Robot> Saving URL ... Wn");
    SaveURL(URL);
    return(SUCCESS)
}
}
```

[그림 2-4] Robot에 의한 문서 수집 모듈

2.2 문서 수집 로봇 에이전트 기술

로봇들은 다음과 같은 많은 유용한 일들을 수행하

는데 사용될 수 있다[2,3,4].

- 통계분석 : 로봇의 목적중의 하나는 웹 서버를 발견하거나 서버의 수를 세는데 사용될 수 있다.
- 유지보수 : 하이퍼텍스트 구조를 유지하는데 주된 어려움중의 하나가 다른 페이지에 대한 레퍼런스가 “죽은 링크(dead links)”가 되는 경우이다. 이러한 경우는 레퍼런스 되어지고 있는 페이지가 옮겨졌거나 없어졌을 때 발생한다.
- 미러링(Mirroring) : 미러링은 FTP 아카이브(archives)를 유지하는 일반적인 테크닉이다. 미러는 FTP에 의해서 순환적으로 디렉토리 트리를 완전히 카피하고 변화된 문서들을 검색한다.
- 리소스 발견 : 아마도 로봇들의 가장 흥미있는 애플리케이션은 리소스 발견의 사용에 있다. 사람이 처리할 수 없는 막대한 양의 정보에 대해 컴퓨터가 하도록 하는 것은 흥미있는 것이고 실제로 몇몇 로봇들은 웹의 많은 부분을 요약하고 있고 서치 엔진을 통해서 이들 결과를 가지고 있는 데이터베이스 접근을 제공한다.
- 복합적인 사용 : 한 로봇은 위의 일중에서 하나 이상의 일을 수행할 수 있다. 로봇의 순회 방법은 다음과 같다.
 - 넓이-우선 순회(breadth-first traversal) : 상위 레벨에서부터 제한된 깊이까지 보다 넓은 그리고 상위 레벨의 문서와 서비스의 집합을 찾는데 좋다.
 - 깊이-우선 순회(depth-first traversal) : 새로운 사이트를 찾는데 보다 좋다. 이 방법은 한 다큐먼트에 여러 레퍼런스가 있을 경우 한 레퍼런스에 대해서 계속 따라가면서 검색하는 방법이다. 서치 엔진에서 자료를 수집하고 인덱싱하는데 로봇 에이전트를 이용할 수 있다. 이것은 로봇의 이용 방법 중 리소스 발견 측면인데, 서치 엔진이 로봇을 동작시켜서 검색된 정보들을 데이터베이스에 저장하고 인덱싱해서 사용자들이 질의를 통해 데이터베이스로부터 정보를 검색할 수 있게 한다. 임의의 웹 문서를 인덱싱하기는 매우 어렵다. 초기의 로봇들은 문서 타이틀과 앵커(anchor)텍스트를 간단히 저장했다. 하지만 새로운 로봇들은 보다 진보된 메카니즘을 사용하고 내용 전체를 고려한다. 이들 방법들은 일반적으로 좋은 수단이고, 모든 웹 페이지들에 자동적으로 적용되어질 수 있다. 그렇지만 손으로 인덱싱 한 것 만큼 효율적일 수 없다. HTML은 <META>를 명시함에 의해 문서에 일

반적인 메타 정보를 덧붙일 수 있는 기능을 제공한다. 예를 들면, “<meta name = "keywords" value = "Ford Car Maintenance">와 같다. 그러나 이 태그 (tag)의 애트리뷰트의 명세된 값에 대해 의미가 정의되어 있지 않기 때문에 그것을 받아들일 것인지 그것의 사용 여부에 대해서는 엄중히 제한된다.

2.3 저장 시스템

저장 시스템은 문서 색인기에 의해 처리된 정보를 저장한다. 저장 시스템은 관계형 데이터베이스인 PostgreSQL DBMS에 의해 관리 되어진다. [표 2-2]는 저장 시스템에 저장된 테이블의 Layout을 보여주고 있다.

주제어	디렉토리	지역	URL
한우	육류	함평	http://www.hampyung.chonnam.kr/~webbbs/bbs-index.html
단감	과일류	무안	http://www.muam.chonnam.kr/special/special4.html
장미	화훼류	강진	http://www.kangjin.chonnam.kr/agriculture/agriculture6.html

[표 2-2] 저장 Table Layout

2.4 검색기

검색기는 로봇에 의해 수집되어 저장된 상품 카달로그 정보를 구매자의 요구에 의해 정보를 추출해 주는 인터페이스이다. 검색 방법에는 주제어의 입력에 의해 질의를 답변해주는 주제어 검색방법과, 상품 정보를 디렉토리화하여 검색방법을 제공하는 디렉토리 검색방법이 있다. 검색엔진과 PostgreSQL과의 인터페이스로는 PHP3.0 (HTML-embedded scripting language)이 사용되어 졌다. PHP는 웹서버에 모듈로 작동한다. [그림 2-5]는 검색 인터페이스 화면이다.



[그림 2-5] 토산넷 인터페이스 화면

2.4.1 주제어 검색 방법

토산넷의 주제어 검색방법은 불리언 검색 모델을 사용하고 있다. 아직까지는 유의어나 동의어를 처리하지는 않고 있다. 주제어 검색은 구매자가 특정 상품목록을 입력하고 검색 버튼을 누르게 되면 관련 상품 목록을 보여주게 된다. 검색절차는 첫째, 데이터베이스의 디렉토리 필드에서 먼저 검사를 한다. 둘째, 첫 번째에서 발견되지 않으면 주제어 필드에서 검사를 하여 원하는 정보를 디스플레이 해 준다. [그림 2-6]은 “공예품”이라는 키워드를 입력하였을 때 보여주는 화면이다.

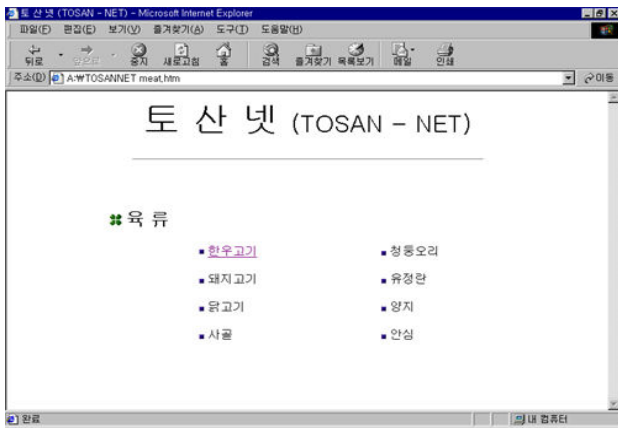


[그림 2-6] “공예품” 주제어 검색 결과

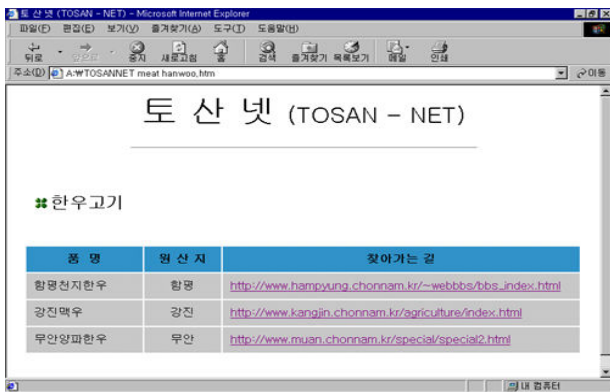
2.4.2 디렉토리 검색 방법

디렉토리 검색 방법은 사용자가 토산품의 분류항목에 의해 검색해 나가는 방법이다. [표 2-1]에서와 같이 “육류, 곡류, 채소류, 과일류, 건어물류, 가공식품류, 공예품, 수산물, 화훼류”의 9가지 대분류를 정해 놓고 구매자가 원하는 항목을 클릭하게 되면 서

브 항목을 디렉토리 색인기가 그때마다 구성하고, 마지막에 데이터베이스에 접근하여 원하는 항목을 보여주게 된다. [그림 2-7,8]는 “육류”중에서 “한우”를 검색하여 추출된 문서 내용을 보여주고 있다.



[그림 2-7] “육류” 디렉토리 화면



[그림 2-8] “한우” 검색 결과

3. 결론

토산넷은 로봇을 이용한 메타 전자 쇼핑몰이다. 대기업이나 또는 메타랜드[5], 현대 백화점[6]과 같은 대형 쇼핑몰들처럼 소비자에게 인지도가 낮은 기업이나 사업자들의 홈페이지에 있는 카탈로그 정보

를 사람이 아닌 로봇이 자료를 수집하여 데이터베이스를 구축하고 서비스를 할 목적으로 자치단체의 특산품과 토산품을 대상으로 시험 구현하였다. 토산넷은 특산품의 정보를 각 자치단체의 웹사이트가 아닌 하나의 사이트에서 서비스를 함으로서 소비자에게 검색 시간을 단축하게 하고, 다양한 정보 서비스를 제공함과 동시에 각 자치단체의 클릭 수를 올릴 수가 있었다. 특히 토산품이라는 특정 제품만을 취급하는 백화점식 쇼핑몰을 구현하므로써 소비자가 특산품이나, 토산품을 구매하기 위해 토산넷으로 오도록 유도할 수가 있었다. 이러한 메타 전자쇼핑몰의 구현은 중소기업과 같은 업체들의 연합으로 구현될 경우 많은 효과가 있으리라 생각된다.

그러나 토산넷을 구현하는 중에 추가되어야 할 몇 가지 문제 해결이 요구되어진다. 첫째는 각 사이트에 있는 전자 카탈로그에 다양한 정보가 있어야 하며, 둘째는 메타 쇼핑몰 사이트에서 구매자들이 구입시 대금을 지불할 수 있는 전자 지불 시스템의 구현이 이루어져야 한다. 이러한 문제만 해결되어지면 이미지가 열악한 기업들의 전자상거래에서 많은 성장 효과가 있으리라고 본다.

참고문헌

- [1] J. Willams, "Bots and Other Internet Beasities", SamsNet. p.257-263, 1996.
- [2] 이은석, “멀티에이전트 기술의 실세계 시스템로의 응용”, 정보과학회지, 제15권, 3호, p17-28, 1997.3.
- [3] Arnold, Ken, and Gosling, James, "The Java Programming Language", Addison-Wesley.
- [4] Fah-Chun Cheong, "Internet Agents: Spiders, Wanderers, Brokers, and Bot", Macmillan Publishing USA, 1995.
- [5] <http://www.metaland.com>
- [6] <http://www.hyundaidept.com>