

링크 맵을 이용한 웹서버 성능 향상 연구

문일형*, 조동섭

*이화여자대학교 컴퓨터학과

e-mail : ilover78@gmail.com, dscho@ewha.ac.kr

The web server performance improvement study using Link-Map

Yil-hyeong Mun*, Dong-sub Cho

*Dept. of Computer Science and Engineering, Ewha Womans University

요 약

최근 대부분의 기업은 인터넷 웹 서비스를 하고 있고, 대부분의 인터넷 사용자들이 여러 종류의 웹 서비스를 제공받고 있다. 하지만 사용자들이 필요로 하는 서비스는 한정되어 있지만 대부분의 기업에서는 여러 가지 서비스를 제공하여 다양한 많은 사용자들의 요구를 수용하고자 한다. 그로 인해 초기 메인 웹 페이지에 많은 링크가 연결된 페이지를 제공하게 된다. 이는 초기 웹페이지의 로딩에 있어 큰 용량의 소스가 원인이 되어 다수의 사용자가 짧은 시간에 접속할 경우 서비스 오류의 원인이 되기도 한다. 또한 대부분의 사용자들이 필요로 하는 몇 개의 링크 기능만을 제공하고 니즈가 적은 링크 기능은 비활성화 함으로써 사용자 요구에 좀 더 빠르게 대응할 수 있도록 링크 맵 프로그램을 이용한 웹 시스템을 제안하고 웹 서버의 성능 향상을 연구한다.

1. 서론

최근 웹 이용자들의 폭발적인 증가와 더불어 인터넷 기업의 다양한 서비스를 위한 초기 웹페이지의 내용이 늘어나고 있다. 초기 웹페이지의 대부분을 차지하는 내용은 각 문자열에 연결된 링크주소이다. 이 링크 주소들은 특히 포털기업 메인 웹 페이지에 다수로 연결되어 있다. 다양한 사용자들의 니즈를 위하여 다양한 서비스에 다양한 링크들이 연결되어 있다. 그러나 이런 다수의 링크 주소들은 메인 웹페이지의 소스의 용량을 늘어나게 하고, 그로 인해 좀 더 많은 사용자가 웹 페이지에 접근하는데 있어 제한을 받을 수 있다. 포털은 물론 홈쇼핑 등의 다수 사용자의 수용과 기존 사용자에 대한 차별화된 링크제공은 응답시간에 민감한 사용자들에게 중요한 이슈가 될 수 있다. 본 논문은 링크정보 및 하이퍼링크에 대한 관련연구와 더불어 링크 맵을 이용한 초기 웹페이지의 렌더링 서비스를 제안하고 구현한다.

2. 관련 연구

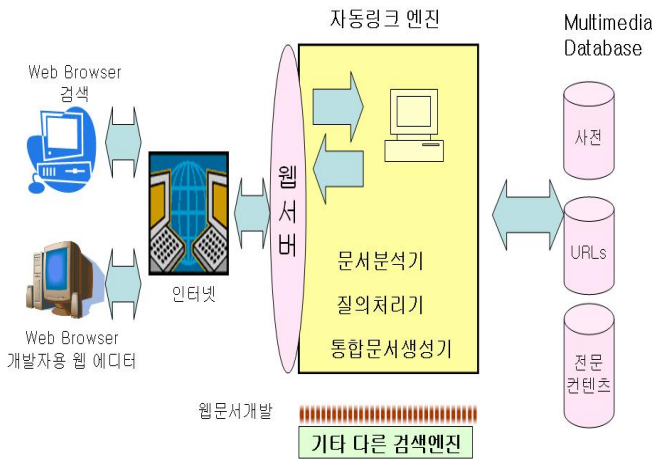
2.1 자동링크지원 시스템(ALSS : Automatic Linking Support System)[3]

ALSS 는 웹 서비스 부분으로, 기존의 문서를 하이퍼 텍스트로 변환하고 연결된 정보를 자동으로 링크시켜 정보를 재사용하며, 기존 정보의 재활용 및 통합문서 작성에 수월하다. 웹 사용자의 전문 콘텐츠 및 전자 사전 및 관련 웹 사이트에 대한 효율적인 검색과 사용자를 위한 맞춤 서비스를 제공한다.

ALSS 의 구성요소는 다음과 같은 기능을 갖는다.

첫째, 원하는 정보의 검색을 위해 기존의 텍스트 직접 기입하여 검색하는 방법에서 포인팅 디바이스를 해당 단어에 위치시킴에 따라 지정된 단어에 대한 정보를 실시간으로 검색할 수 있도록 한다. 둘째, 사용자가 하이퍼미디어 문서를 작성할 때, 사용자가 작성한 텍스트 문서에 포함된 단어들에 대한 정보를 데이터베이스로부터 검색 후 추출하여 텍스트 문서의 단어에 자동링크시켜 통합된 하이퍼미디어 문서를 생성할 수 있도록 한다. 마지막으로 통합된 하이퍼미디어 문서를 자유로이 사용자의 의도에 맞도록 재편집 및 실시간 검색을 제공할 수 있도록 한다. 이를 위해 ALSS 는 클라이언트/서버 구조로서 인터넷 기반에서 수행된다. 서버 측은 전문 콘텐츠, 용어사전 및 관련 웹 사이트를 제공하기 위한 데이터베이스 시스템과 웹 문서 작성을 웹 에디터와 연결되어 통합된 하이퍼미디어 생성 및 편집을 지원하는 자동링크엔진으로 구성된다. 그리고 클라이언트 측은 웹 에디터와 웹 도우미로서 서버측과 연동하여 하이퍼링크된 통합문서 생성과 검색 서비스를 요청한다.

클라이언트인 웹에디터와 웹도우미, 서버의 역할을 하는 자동링크엔진, 그리고 여러 정보를 포함하고 있는 데이터베이스 시스템의 3 계층의 구조로 구성될 수 있다.

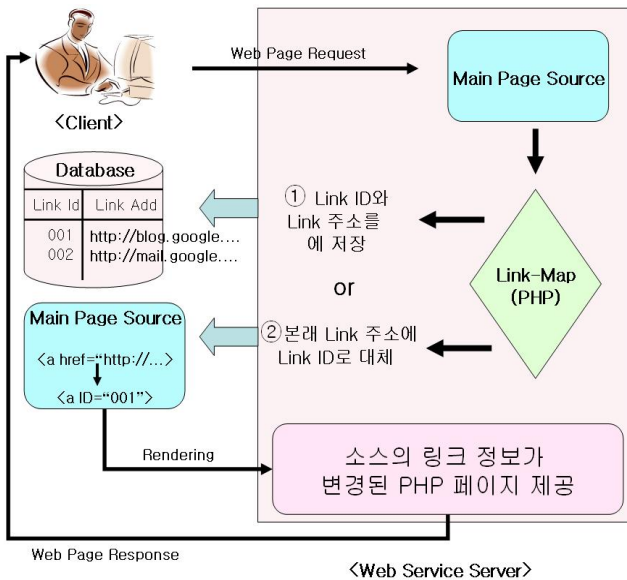


(그림 1) ALSS의 전체구조

3. 링크 맵 웹 시스템

3.1 링크 맵 웹 시스템 설계

링크 맵 웹 시스템은 위 관련연구에서 언급된 ALSS의 시스템과 달리 현재 너무 많은 하이퍼링크 정보들이 구현된 웹 페이지의 소스를 효율적으로 줄이고자 웹 이용자가 이용하는 부분의 문자열에만 하이퍼링크 기능을 구현하고 그 외의 부분은 하이퍼링크 기능을 구현하지 않는다. 이는 초기 웹페이지의 빠르고 효율적인 로딩을 할 수 있도록 한다.



(그림 1) 링크 맵 웹 시스템 구조

위 그림은 링크 맵 웹 시스템의 구조를 나타내고 있다. 링크 맵의 기능은 우선 사용자가 웹페이지를 초기에 요구했을 때, 시스템이 가장 처음 구동 됐을 경우 링크 맵은 웹페이지의 링크 주소를 데이터베이스에 차례대로 저장하게 된다. 이때 각 링크 주소를 구분하기 위해 링크 ID와 함께 데이터 베이스에 저장하게 된다. 이는 웹페이지를 다시 작성할 때 필요

한 링크만 가져오기 위한 사전 작업이기도 하다. 이 작업은 최초 페이지를 변환할 때 외에는 작업하지 않아도 된다.

저장된 데이터베이스 정보를 가지고 본래 웹페이지의 링크들 중에서 사용자가 필요로 하지 않는 문자열에 대한 링크는 ID 정보로 대체함으로써 비활성화하도록 한다. 이는 웹페이지 이용자들이 웹페이지 초기 화면의 모든 서비스를 차례대로 이용하는 것이 아니라 사용자가 이용하는 부분이 한정되어 있기 때문에 사용자가 주로 이용하는 문자열에 대한 하이퍼링크 기능만 활성화 하고 그 외의 문자열의 하이퍼링크 기능은 비활성화 함으로써 기본 웹페이지의 소스 크기를 줄일 수 있게 된다. 메인 페이지의 <a href>태그를 라벨태그로 바꿈과 동시에 ID는 데이터베이스의 ID로 대체함으로써 다른 사용자를 위한 페이지를 다시 구동할 경우 데이터베이스의 본래 링크 정보를 손쉽게 찾을 수 있게 된다.

3.2 링크 맵 웹 시스템 구현

링크 맵 웹 시스템의 구현은 일반 검색이 가능한 웹 페이지를 대상으로 링크 맵 웹 시스템이 구동되어 소스의 링크 정보가 변경된 PHP 파일을 작성하는 것이다.

<표 1> 링크 맵 웹시스템 구현을 위한 환경 설정

Web Server	Apache
Language	PHP
Database	MySQL
Test Web Page	http://scholar.google.co.kr

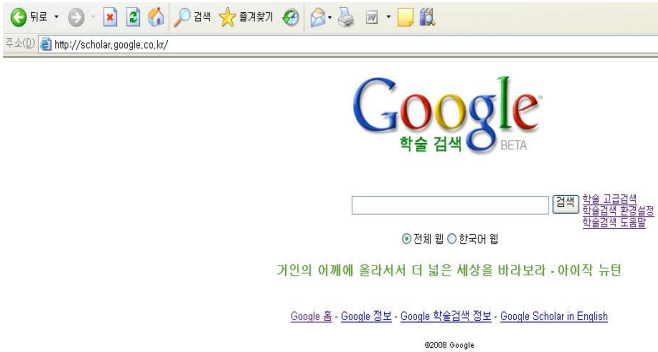
위 표와 같은 시스템 환경을 기본으로 링크 맵 기능의 PHP 프로그램을 작성하였다. 이 PHP 파일이 일반 링크 정보가 여러 개의 링크 정보가 들어간 소스를 필요한 링크 정보가 구동되게 하는 렌더링 작업을 하게 된다. 렌더링 작업으로 만들어진 사용자가 필요로 하는 링크만 활성화 하도록 한다.

```

22 for($i=0;$i<count($arr);$i++){
23     $arrs=str_replace($search,$replaces,$arr[$i]);
24     $arrs=str_replace($search,$replaces,$arr[$i]);
25     fwrite($fp, $arrs);
26 }
27
28 //fwrite($fp, $arrs);
29 $db=new mysqli('localhost','root','','linkmap');
30 if(mysqli_connect_errno())
31 {
32     echo 'Error: Could not connect to database. Please try again later.';
33     exit;
34 }
35
36 $query = "insert into linkmap values
37 ('".$linkID."','".$linkAdd."")";
38 $result = $db->query($query);
39 $db->close();
40
41 fclose($fp);
42
43 $fp = fopen("$DOCUMENT_ROOT/phantest/chapter02/sch.txt", 'rb');
44
45 //string = "<a href=
46 if (!$fp)
47 {
48     echo '<p><strong>No orders pending.
49     .Please try again later.</strong><p>';
50     exit;
51 }
52
53
54 while (!feof($fp))
55 {
56     $order= fgets($fp, 999);
57     echo $order."<br>";
58 }
59
60
61 fclose($fp);
62
    
```

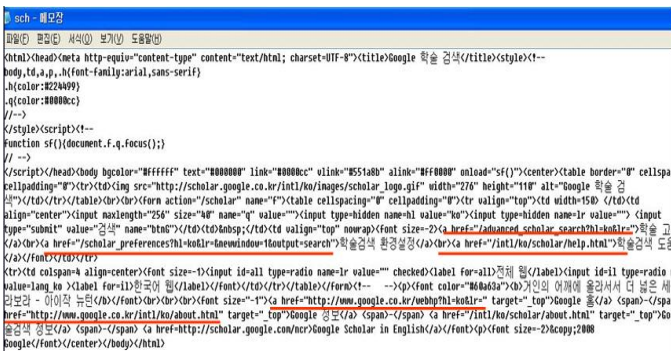
(그림 2) 링크 맵 PHP 프로그램

위 PHP 파일을 통해 웹 페이지 소스가 렌더링 작업을 통해 사용자가 필요로 하는 링크만 남게 되는 구현은 다음 그림을 보면 알 수 있다. 테스트 페이지는 구글 학술검색 페이지의 소스를 링크 맵으로 렌더링 한다. 먼저 렌더링 전의 <http://scholar.google.co.kr>의 웹 페이지를 링크 맵으로 렌더링하여 php 파일로 변환하고 그 후 한번 더 사용자가 사용하지 않는 링크는 비활성화하도록 한 번 더 렌더링 한다.



(그림 3) <http://scholar.google.co.kr>의 메인 페이지

위의 메인 페이지를 갖는 소스를 링크 맵 php 프로그램으로 1 차로 렌더링 한다. 위 그림 3의 페이지의 소스를 보면 ``로 연결된 부분들이 링크 기능하는 부분이다.



(그림 4) 1 차 렌더링 한 웹 페이지 소스와 하이퍼링크 부분

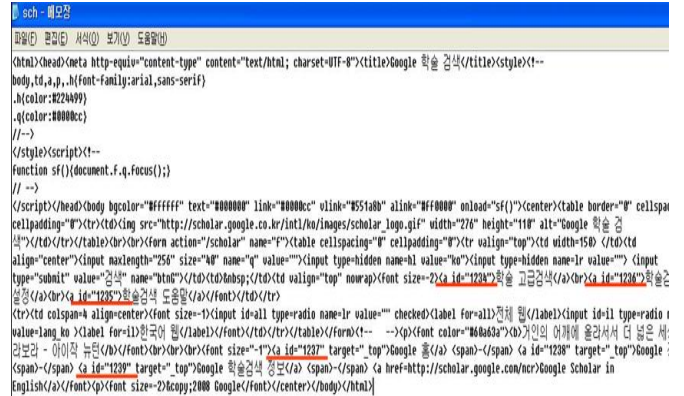
<http://localhost/phetest/chapter02/linkmap.php>

Link Map Test - Mun Yilhyeong



(그림 5) 1 차 렌더링 한 메인페이지

위의 웹 페이지는 링크 맵으로 한번 렌더링 한 후 보여지는 웹 페이지다. 위 그림에 표시된 사용자가 사용하지 않는 링크에 대해서 변환을 한다. 링크의 기능이 아직 활성화 된 상태로서 linkmap.php 파일로 변환만 된 상태이다. 소스 부분의 링크 정보가 ID 태그로 변환함과 동시에 데이터베이스의 ID 정보가 들어가면서 그림 4의 하이퍼링크는 비활성화 하게 된다. 그 동작을 구현한 후 결과는 다음과 같다.



(그림 6) 2 차 렌더링 한 웹 페이지 소스와 하이퍼링크 부분

<http://localhost/phetest/chapter02/linkmap.php>

Link Map Test - Mun Yilhyeong



(그림 7) 2 차 렌더링 한 메인페이지

위 그림 6의 소스를 보면 2 차로 링크 정보를 ID 태그로 변경하고 데이터베이스의 ID 로 변환한 소스를 보여준다. 1 차 렌더링한 소스의 링크정보와 비교해 볼 때 소스의 길이가 줄어든 것을 확인 할 수 있다. 그림 7을 보면 사용자가 바로 검색을 할 경우, 그 외의 링크 기능을 비활성화하고 서비스의 여부만을 보여 주게 된다. 사용자가 검색만 할 경우 굳이 링크 기능을 모두 활성화 할 필요는 없기 때문에 전혀 검색하는데 있어서 문제가 발생하지 않는다.

검색 사용자가 필요로 하는 검색부분만을 활성화 함으로써 사용자의 메인 웹페이지의 로딩을 빠르게 하고 동시에 다수의 사용자가 더 많이 접근 할 수 있게 된다.

4. 결론 및 향후 연구

현재 많은 포털 기업과 더불어 쇼핑몰 또는 일반 기업의 웹 사이트에서 많은 서비스를 수 많은 사용자에게 제공하고자 하는 의도로 인해 첫 메인 웹 페이지에 다양한 링크를 연결 하고 있다. 이는 다양한 사용자의 수용은 가능 할 수 있으나 고정적으로 일정 서비스만을 사용하는 사용자에게는 메인 페이지의 소스가 커지게 됨으로 접근자가 단시간 많을 경우 오버헤드를 일으킬 수 있다. 이러한 문제를 좀 더 효율적으로 관리하기 위해서 링크 맵 프로그램을 통해 사용자가 필요로 하는 링크만을 활성화 함으로써 웹 페이지 소스의 크기를 줄이고 동시에 좀 더 많은 사용자의 접근이 가능하다. 이는 일정한 링크만을 접근하는 고정 사용자에게 더욱 유용한 기능이라고 본다.

추후 각각의 접근 사용자 또는 일정 그룹의 사용자들을 위한 링크를 결정하는 연구가 필요하다. 각각의 접근 사용자를 위한 링크만이 제공되는 경우는 사용자 입장에서는 매우 유용할 것이다. 그러나 현재 수많은 사용자의 링크 접근 패턴을 찾아내고 그 정보를 저장하는데 적지 않은 어려움이 있을 수 있다. 그리하여 접근 패턴이 유사한 사용자들을 클러스터 한 후 링크를 정보를 제공하는 것도 연구되어야 할 것이다.

참고문헌

- [1] Yu Bo, Qi Luo. "Personalized Web Information Recommendation Algorithm Based on Support Vector machine", International Conference on Intelligent Pervasive Computing, pp487-pp490, 2007
- [2] Lan Nie, Brian D. Davison, Xiaoguang Qi. "Topical Link Analysis for Web Search", SIGIR04, August 6-11, pp91-pp98, 2006,
- [3] J.M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604-632, 1999
- [4] 이원중, 정은재, 주수종, 이승용. "웹 통합문서의 효율적 생성과 검색을 위한 자동링크지원 시스템의 설계 및 구축", 정보처리학회논문지 A 제 10 권-A 권 제 2 호, 2003
- [5] 염미령, 홍일구. "작업의 잔여량을 이용한 웹서버 로드 밸런싱", 정보처리학회 추계학술발표논문집, 제 8 권 제 2 호, 2001
- [6] 이종섭, 최영근. "인터넷 스크립팅 언어의 동향 및 응용에 관한 연구", 정보처리학회 논문지 제 6 권 제 11 호, 1999
- [7] Google scholar, <http://scholar.google.co.kr>