

Sitemarking: 모바일 사용자를 위한 북마크 및 추천기법

신소연⁰ 황인준
아주대학교 정보통신전문대학원 정보통신공학과
{seaaes3⁰, ehwang}@ajou.ac.kr

Sitemarking: Bookmark and recommendation for mobile user

Sohryun Shin⁰ Eenjun Hwang
Graduate School of Information and Communication, Ajou University

요약

무선 단말기 보급의 증가로 PDA 또는 휴대폰 등을 통한 인터넷 접속이 증가하고 있다. 하지만 현재 대부분의 웹 사이트는 페이지 크기나 용량이 데스크톱을 위해 최적화 되어 있기 때문에, 무선 단말기를 이용한 원활한 인터넷 서핑이 어렵다. 이런 문제에 대한 해결책으로 개인화된 북마크를 통해 원하는 서비스에 바로 접근할 수 있는 방법을 제안한다. 본 논문에서 제안한 기법은 기존 페이지 중심의 북마크를 개선하여 웹 사이트 별로 일련의 페이지들을 북마크하여 무선 단말기를 통해 제공함으로써, 웹 사이트의 구조를 한눈에 쉽게 파악하고 원하는 페이지에 더욱 빠르게 접근할 수 있는 방안을 제시한다. 또한 사용자들의 선호도 프로파일을 이용하여 유사한 관심분야를 갖는 사용자들을 그룹화한 후 같은 그룹내의 사람들이 생성한 북마크를 웹 사이트 접근시 추천해주는 기법을 제안한다.

1. 서론

최근 무선 인터넷 환경의 급속한 성장과 무선 단말기 보급의 증가로 인해, PDA(Personal Digital Assistant)나 휴대폰(Cell Phone) 등을 통한 인터넷 접속이 증가하고 있다. 인터넷 통계 사이트인 Comscore Network의 최근 기사에 따르면, 현재 미국에서 판매되고 있는 휴대폰의 95% 이상이 웹에 접속할 수 있으며, 이미 990만명 이상의 사람들이 PDA나 휴대폰을 이용해서 인터넷에 접속하고 있다[1].

그러나 이런 무선 인터넷 접속의 폭발적인 증가에도 불구하고, 소수의 웹 사이트만이 무선 단말기 전용 웹 콘텐츠를 제공하고 있으며, 그 외의 대부분의 웹 사이트들은 데스크톱에 최적화 되어있다. 그러므로 상대적으로 작은 화면, 낮은 대역폭 및 제한적인 입력장치를 가진 무선 단말기를 이용한 원활한 인터넷 서핑이 어렵다. 이런 문제로 인해 무선 단말기에 웹 콘텐츠를 요약하여 재구성해서 보여주거나[2,3] 무선 단말기를 위한 페이지 단위의 북마크를 이용하는 새로운 분야의 연구가 행해졌다. 하지만 이런 연구들은 웹 콘텐츠 변환 과정에 과도한 부하가 발생하여 실시간 변환이 어렵거나, 북마크 생성과정에서 링크의 구조가 새롭게 바뀌어 사용자의 적응에 어려움을 주는 등의 단점을 가지고 있다.

본 논문에서는 이런 단점을 극복하기 위해서 사용자들이 선호도별로 저장한 북마크를 재구성하여 무선 단말기에 적합하게 제공함과 동시에 유사한 관심분야를 가진 사용자에게 웹 사이트의 북마크를 추천해주는 Sitemarking 기법을 제안한다. 이는 사용자가 웹 사이트상에서 이동했던 링크 경로를 텍스트 형식의 트리 리스트로 제공함으로써 사용자가 자주 이용하는 웹 콘텐츠에 쉽게 접근할 수 있고 구조를 한눈에 알아 볼 수 있다. 또한 데스크톱에서 사용자가 사전에 북마크 하지 못했

던 사이트를 무선 단말기에서 요청시, 유사 관심그룹에 속해 있는 다른 사용자의 북마크를 유사도가 높은 순서로 추천해 준다.

본 논문의 구성은 다음과 같다. 2장에서는 무선 단말기를 위한 웹 페이지 요약 및 개선된 북마크 시스템에 관련된 연구들을 논의한다. 3장에서는 본 논문에서 제안하는 북마크 생성 기법과 Sitemarking 시스템의 구조를 기술한다. 4장에서는 사용자들을 관심분야에 따라 그룹화 하는 기법 및 이에 따른 북마크 추천 시스템에 대해 기술한다. 5장에서는 서비스 시나리오를 살펴보고 끝으로 6장에서는 결론 및 향후 과제를 제시한다.

2. 관련연구

무선 단말기를 위한 웹 콘텐츠 요약 및 재구성에 관한 연구인 Digstor[2]는 중요하지 않는 부분의 내용을 생략하거나 내용을 다음 페이지의 링크로 연결하는 등 화면의 구성을 바꾸었다. 유사한 연구인 Power Browser[3]는 HTTP proxy를 이용해 동적으로 웹 페이지의 요약된 뷰를 생성하여 사용자에게 제공해준다. 하지만 위 연구들은 사용자의 의사반영이 없이 일방적으로 내용을 요약하며, 이 과정에서 본래 웹 페이지들간 링크구조가 재구성 되어 전체 웹 사이트의 구조를 파악하기 어려우며 사용자가 새로운 링크구조에 적응하기 위해서는 다소 시간이 소요된다. 또한 요약된 페이지라 하더라도 타겟 페이지에 접근하기 위해서 중간 페이지들을 순차적으로 다운로드 해야 하기 때문에, 낮은 대역폭과 제한된 입력장치로 타겟 페이지까지 접근하기 위해선 많은 인내가 필요하다. 이런 단점을 극복하기 위해 무선 단말기를 위한 새로운 북마크 시스템에 대한 연구들이 진행되어왔다.

WebView[4]는 사용자가 원하는 서비스로 이동하기까지 요

구되는 경로 및 인증 정보를 저장함으로써 SmartBookmark를 생성한다. 그러나 SmartBookmark는 사용자간 공유가 불가능하며 사전에 북마크 하지 않은 사이트들에는 적용시킬 수 없다. 다른 북마크 관련 연구인 Tsuyoshi[5]는 페이지 내의 관심 있는 문장에 북마크 함으로써 재접근 시 해당 콘텐츠의 위치로 빠르게 이동할 수 있도록 한다.

3. Sitemarking 시스템

Sitemarking 시스템은 기존의 데스크톱에서 제공되고 있는 페이지별 북마크 방식을 개선한 웹 사이트내 일련의 페이지들을 순차적으로 북마크 해주는 서비스이다. 현재 데스크톱에서는 그림 1-(a)와 같이 페이지 단위의 북마크를 제공한다. 페이지 단위 북마크는 사용자가 웹 사이트의 링크를 따라가면서 원하는 웹 페이지가 나오면 링크의 마지막 페이지를 북마크를 하는 것을 말한다.

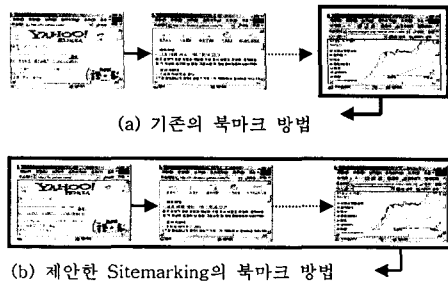


그림 1. Sitemarking과 기존 북마크의 차이

이런 기존의 북마크 방법은 인터넷에 접속하자마자 해당 웹 페이지를 볼 수 있지만, 그 페이지의 상위링크나 하위링크의 웹 페이지에 접근하기 위해서는 첫 페이지부터 계속 링크를 따라가야 하는 단점이 있다. 또한 무선 단말기는 데스크톱에 비해서 대역폭이 낮기 때문에, 일반 데스크톱에 최적화된 웹 사이트를 불러오는 것은 비효율적이다. 이런 단점을 해결하기 위해서 본 논문에서는 그림 1-(b)와 같이 사용자가 관심 있는 웹 사이트내 일련의 웹 페이지들을 북마크하여 웹 사이트내에서 원하는 웹 페이지를 신속히 접근할 수 있는 기법을 제안한다. 그림 2는 시스템의 전체적인 구조를 보여준다.

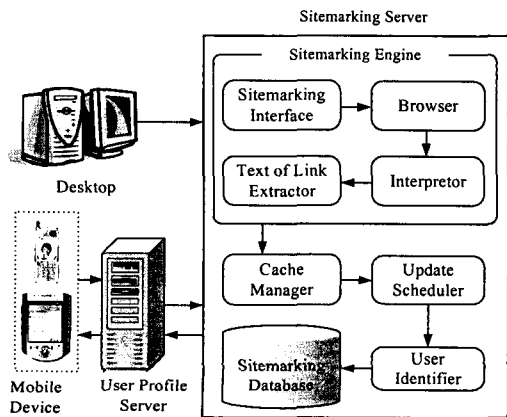


그림 2. 시스템의 전체적인 구조

Sitemarking Server는 다음과 같은 구성요소를 포함하며 상호 동작을 처리한다.

- ① **Sitemarking Database:** 사용자의 개인정보와 북마크 정보 및 그룹별 정보를 포함하고 있다.
- ② **User Identifier:** 서버에 있는 데이터베이스에 북마크 정보를 저장할 때, 사용자의 인증 확인 등을 한다.
- ③ **Cache Manager:** 웹 사이트 경로 등을 저장하는 북마크 정보를 캐쉬한다.
- ④ **Update Scheduler:** 주기적으로 북마크의 정보를 갱신한다. 예를 들어 갱신(update) 기간을 한 달로 설정하면, 설정한 기간이 지난 후 기존 정보를 새로운 북마크 정보로 갱신한다. 이런 기능이 필요한 이유는 한 사이트에서 북마크를 지정할 수 있는 경우(포털 사이트인 야후의 경우 날씨, 매일, 주식, 뉴스, 예약과 같은 서비스 종류의 수)의 수를 5가지로 한정시켰기 때문이다.
- ⑤ **Sitemarking Engine:** HTML 문서를 검색 및 파싱해서 북마크한다. 또한 Text of Link Extractor는 사용자가 웹 사이트에서 원하는 정보를 따라갈 때, 자바 API를 이용해서 링크되는 부분의 텍스트를 추출한다.

4. 그룹화(Grouping)와 추천(Recommendation) 시스템

본 논문에서 제안하는 Sitemarking 시스템은 웹 사이트내 링크들의 북마크와 유사한 관심분야를 가진 그룹 내의 사용자가 만든 북마크를 추천해 주는 것으로 구성된다. 여기서 그룹화란 유사한 성향의 사용자들을 일괄처리하기 위한 것이고, 추천 시스템은 사용자가 사전에 북마크 하지 못했던 사이트 요청시, 유사 관심그룹에 속해 있는 다른 사용자의 북마크를 유사도 순으로 추천해 주는 것이다.

먼저 사용자들의 그룹을 나누기 위해서 CC/PP(Composite Capability Preference Profiles)[6]로 작성된 장치 프로파일, 사용자 프로파일을 이용한다. CC/PP는 Content Negotiation에 관련된 표준으로 단말기 및 사용자의 성향을 기술한다. 이는 무선 단말기 등의 전송 환경을 인식한 후, 인지된 전송 환경에 적합한 형태로 웹 콘텐츠를 변환한다. 본 논문에서는 그림 3과 같이 제정의한 RDF를 사용하였다.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:prf="http://adtl.ajou.ac.kr/schema#"
>
<rdf:Description
rdf:about="http://adtl.ajou.ac.kr/profile#MyProfile">
<ccpp:component>
<rdf:Description
rdf:about="http://adtl.ajou.ac.kr/TerminalHardware">
.....
<prf:vendor>Samsung</prf:vendor>
<prf:ScreenSize>240*320</prf:ScreenSize>
.....
<rdf:Description
rdf:about="http://adtl.ajou.ac.kr/TerminalBrowser">
.....
<prf:Computer>Yes</prf:Computer>
<prf:Beauty>Yes</prf:Beauty>
<prf:News>No</prf:News>
<prf:Economy>Yes</prf:Economy>
.....
</rdf:RDF>
```

그림 3. CC/PP파일의 일부분

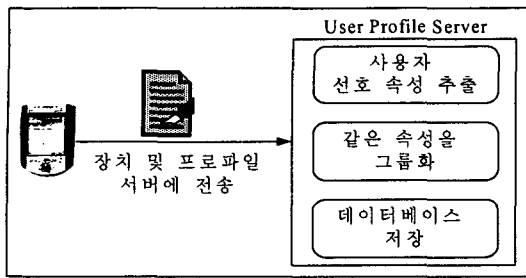
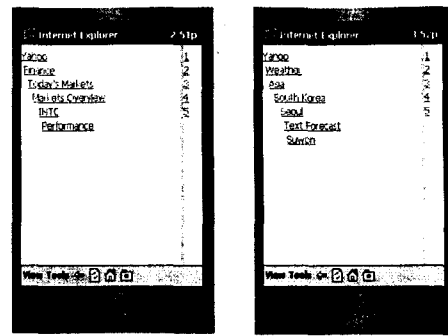


그림 4. 사용자 속성 해석 및 그룹화 과정

작성된 프로파일을 그룹으로 묶는 과정은 그림 4와 같다. 사용자가 작성한 CC/PP파일은 User Profile Server에 전송한다. 그 다음 단계는 사용자 관심 속성을 추출하는 단계이다. 사용자의 관심 속성은 정치, 경제, 컴퓨터, 여행, 교육 각 4개씩 총 20개의 항목을 체크박스 형식으로 정해 놓은 뒤, 사용자가 관심있는 항목에 체크를 한다. 사용자의 관심 속성은 DOM을 이용하여, 미리 정의해 놓은 태그 중에서 속성값이 "Yes"인 것들을 추출한다. 이 중에서 같은 속성들을 가진 사용자들을 한 그룹으로 만든다. 이렇게 각 분야의 항목들이 완전히 같은 사용자 그룹을 유사도 100%로 한다. 하지만 총 20개의 항목에서 한 개의 항목만 틀려도 완전히 다른 그룹으로 분류했기 때문에, 어떤 그룹에는 사용자가 한명만 있을 수도 있는데, 이런 경우는 북마크를 추천해 줄 수 없다. 따라서 이러한 경우를 보완하기 위해 유사도가 70%이상이면 같은 그룹에 포함시키기로 한다. 예를 들어 컴퓨터 분야 중 2개, 경제 분야 중 2개로서 분야 및 선택 항목은 같지만, 분야별로 하나씩만 틀린 경우는 유사도 75%로써 같은 그룹으로 포함한다. 이렇게 저장된 정보가 Sitemarking Database와 연동해서 그룹별 정보를 만든다. 그룹별 테이블들은 User Profile Server로부터 온 정보를 Sitemarking Database의 테이블에 입력한다. 사용자가 북마크 하지 않는 정보를 요청 시, 검색해서 제일 유사도가 높은 순으로 추천하도록 한다.

5. 서비스 시나리오

본 논문에서 최종적으로 제공하고자 하는 서비스의 시나리오를 작성하면 다음과 같다. 사용자가 데스크톱을 통해 Sitemarking Server에 접근한 후, 웹 사이트를 탐색하면서 북마크를 한다. 그 후 무선 단말기에서 북마크한 사이트를 요청하면, 사용자가 작성한 북마크를 링크 형식으로 텍스트로 보여준다. 예를 들어 사용자가 야후 사이트에서 Finance, Today's Market 등의 링크를 계속 클릭하면서 웹 사이트를 북마크 한 경우, 무선 단말기에서 요청하면 그림 5-(a)와 같은 서비스가 제공된다. 또한 화면 우측의 숫자는 웹 사이트내 북마크할 수 있는 개수를 의미한다. 즉 야후와 같이 포털 사이트에서는 날씨, 주식 등의 이용서비스가 다양해서 한 웹 사이트에서 최대 5개까지 북마크 할 수 있도록 설정했다. 그림 5-(a)에서 화면 우측의 링크 '2' 를 클릭하면, 그림 5-(b)와 같은 야후 사이트내 날씨에 대해 북마크한 정보가 제공된다. 그리고 이러한 정보는 Update Scheduler에 의해 주기적으로 자주 이용하는 북마크 정보로 갱신된다.



(a) 증권부분 북마크 (b) 날씨부분 북마크
그림 5. 제공되는 서비스 시나리오

6. 결론

인터넷 사용이 가능한 무선 단말기들의 보급이 크게 증가하면서 무선 인터넷 사용자의 수가 지속적으로 증가하고 있다. 그러나 대부분의 웹 사이트들은 데스크톱 환경에 최적화되어 있어 무선 단말기를 통한 웹 접근이 쉽지 않다. 이런 문제의 해결책으로 무선 단말기를 위한 페이지 단위의 북마크를 이용하는 새로운 분야의 연구가 행해졌다.

본 논문에서는 웹 사이트 단위의 북마크와 추천 사이트에 초점을 둔 Sitemarking 기법을 제안하였다. 개인에 최적화된 북마크 시스템을 사용함으로써 무선 단말기의 제약을 완화함과 동시에 사용자가 관심 있는 웹 콘텐츠를 더욱 쉽게 이용할 수 있도록 하였다. 또한 자신이 설정해 놓지 않은 북마크를 요청했을 경우, 같은 그룹내의 사용자가 작성해 놓은 북마크를 추천받을 수 있다. 이로 인해 개인화 시스템에서 예측을 벗어났을 경우의 대안책을 제시해주기 때문에 기존의 시스템에 비해서 효율적이다.

참고문헌

- [1] <http://www.comscore.com/> --comScore Networks, Inc.
- [2] Timothy W. Bickmore, B. N. Schilit, "Digester: Device-independent access to the World Wide Web", In Proceedings of the 5th International World Wide Web Conference, 1996
- [3] Orkut Buyukkokten, Hector Garcia-Molina, Andreas Paepcke, Terry Winograd, "Power Browser: Efficient Web Browsing for PDAs", Digital Libraries Lab, Stanford University, 2000
- [4] Juliana Freire, Bharat Kumar, D. Liewen, "WebView: Accessing Personalized Web content and Services, In Proceedings of the 10th International World Wide Web Conference, 2001
- [5] Tsuyoshi E, Seiji I, and Teruhisa M., "Fast Web by Using Updated Content Extraction and a Bookmark Facility", the 4th Int'l ACM Conference on Assistive Technologies, 2000
- [6] <http://www.w3.org/Mobile/CCPP/>