

# 모바일 에이전트 환경에서 사이트 단위의 개인화된 북마크 제공 기법

신소연<sup>0</sup> 황인준  
아주대학교 정보통신전문대학원 정보통신공학과  
{seaaes3<sup>0</sup>, ehwang}@ajou.ac.kr

## A scheme for personalized bookmark services in a mobile agent environment

Sohryun Shin<sup>0</sup> 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 데스크톱에 최적화 되어 있기 때문에, 무선 단말기를 이용한 원활한 인터넷 서핑이 어렵다. 본 논문에서는 이러한 문제점을 해결하기 위해서 개인화된 북마크를 통해 원하는 서비스에 바로 접근할 수 있는 방법을 제안한다. 제안한 기법은 기존 페이지 중심의 북마크를 개선하여 웹 사이트별로 일련의 페이지들을 북마크하여 무선 단말기를 통해 제공함으로써, 웹 사이트의 구조를 한눈에 쉽게 파악하고 원하는 페이지에 더욱 빠르게 접근할 수 있는 방안을 제시한다. 또한 모바일 에이전트환경에서 사용자의 선호도 프로파일과 유사한 사용자들이 해놓은 사이트 단위의 북마크를 추출한 후, 사용자에게 제공해주는 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 데스크톱 환경에 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 수행되었다. 하지만 이런 연구들은 웹 콘텐츠 변환 과정에 과도한 부하가 발생하여 실시간 변환이 어렵거나, 북마크 생성과정에서 링크의 구조가 새롭게 바뀌어 사용자의 적용에 어려움을 주는 등의 단점을 가지고 있다.

본 논문에서는 이런 단점을 극복하기 위해 페이지단위의 북마크가 아닌 사이트 단위의 북마크를 하며, 에이전트(agent)를 이용하여 자신의 프록시에 없는 사이트의 북마크를 다른 프록시에서 가져와 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 있으며, 상위나 하위링크로 이동이 가능하다. 또한 사용자가

설정해놓은 프록시에 요청한 사이트 단위의 북마크가 없으면, 에이전트가 다른 프록시에 접근해서 유사한 성향을 가진 사용자가 해 놓은 북마크 정보를 가져와 보여준다.

본 논문의 구성은 다음과 같다. 2장에서는 무선 단말기를 위한 웹 페이지 요약, 모바일 에이전트에 관한 연구 및 무선 개선된 북마크 시스템에 관련된 연구들을 논의한다. 3장에서는 본 논문에서 제안하는 북마크 생성기법과 *Sitemarking* 기법의 구조를 기술한다. 4장에서는 의미있는 북마크 정보를 얻기 위한 사용자 웹 검색 순서 및 성향을 추출하는 기법에 대해서 논의한다. 5장에서는 다양한 무선단말기에서 이용가능하도록 정적 프로파일 및 추천기법에 대해서 설명하고 끝으로 6장에서는 결론을 제시한다.

### 2. 관련연구

무선 단말기를 위한 웹 콘텐츠 요약 및 재구성에 관한 연구인 *Digestor*[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 어렵고 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 위해서 많은 시간이 필요하다. 이런 단점을 극복하기 위해 무선 단말기를 위한 새로운 북마크 시스템에 대한 연구들이 진행되어왔다.

모바일 에이전트 관한 연구인 Debbie[4]는 개인화된 정보 검색 시스템을 제공하기 위해서 에이전트 개념을 이용했다. 이것은 개인화된 에이전트 시스템은 사용자의 인터넷 이용과 문맥정보를 이용해서, 사용자들이 필요할 때 제공해주는 기법에 관한 것이다.

### 3. Sitemarking 기법

Sitemarking 기법은 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 북마크(page unit of bookmark)를 제공한다. 페이지 단위 북마크는 사용자가 웹 사이트의 링크를 따라가면서 원하는 웹 페이지가 나오면 링크의 마지막 페이지를 북마크 하는 것을 말한다.

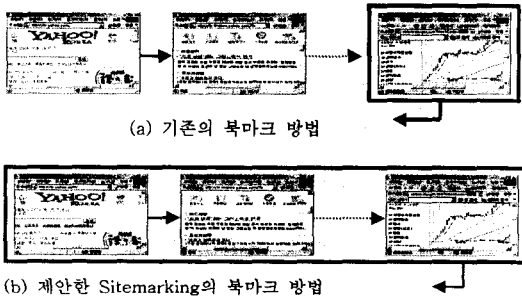


그림 1. Sitemarking과 기존 북마크의 차이

이런 기존의 북마크 방법은 인터넷에 접속하자마자 해당 웹 페이지를 볼 수 있지만, 그 페이지의 상위링크나 하위링크의 웹 페이지에 접근하기 위해서는 새로운 URL을 입력한 후, 계속 링크를 따라가야 하는 단점이 있다. 이런 단점을 해결하기 위해, 본 논문에서는 사이트 단위의 북마크(site unit of bookmark)방식을 제안하며, 이러한 북마크를 Sitemarking links로 부른다. 또한 자신이 설정해 놓은 프록시에 북마크 저보가 없으면, 모바일 에이전트를 통해 다른 프록시에서 정보를 가져와 제공해준다.

### 4. 사용자 웹 검색 순서 및 성향 추출 기법

본 논문에서 제안하는 Sitemarking links를 제공하기 위해서는 사용자의 정확한 웹 탐색순서와 흥미분야를 추출해야 된다. 또한 사용자가 탐색하는 모든 탐색 경로를 북마크 할순 없기 때문에, 이 중에서 의미있는 탐색 경로만을 추출하는 모델이 필요하다.

#### 4.1 사용자의 탐색 패턴 순서 및 관심 영역 추출

웹 사이트내에서 사용자의 의미있는 접근 경로를 추출하기 위해 웹 사이트에서 사용자의 검색 패턴을 알아야한다. 따라서 본 논문에서는 웹 환경에서 사용자의 검색 패턴을 찾기 위해 순회패턴(Traversal Pattern)[5] 알고리즘을 이용해서 사용자의 웹 사이트 검색 순서를 분석한다.

웹 사이트의 페이지들이 서로 연결되어 있다고 가정할 때, 사용자는 자신이 원하는 정보를 얻기 위해 웹 사이트의 링크를 따라 앞이나 뒤로 페이지를 검색하는 경향이 있다. 그 결과로 어떤 페이지의 방문이 내용보다는 위치 때문에 방문될 가능성이 있다.

예를 들어 웹 사이트를 하나의 트리 구조(tree structure)로

생각해 볼 경우, 형제 노드위치에 있는 페이지를 방문하기 위해 사용자는 새로운 URL을 이용하는 대신에 보통 역방향 아이콘(backward icon)을 사용하여 형제 노드위치의 페이지를 다시 방문한다. 결과적으로 웹 사이트내에서 사용자의 기록을 가지고 있는 로그화일에서 의미있는 사용자 접근 패턴을 추출하기 위해서는 이러한 역방향 검색의 영향을 고려하여 실제로 관심있는 접근 패턴을 발견해야한다. 이러한 관점에서 본 논문은, 역방향에 대한 참조는 내용을 검색하기 보다는 검색의 편리성을 위해 이루어진다고 가정한다. 그러므로, 같은 사용자 접근에 의한 역방향 참조는 이전에 방문한 객체를 재방문하는 것을 의미하므로, 역방향 참조가 일어나면, 순방향 참조 경로는 종결된다. 이렇게 얻어지는 순방향 참조 경로를 최대 순방향 참조(Maximal Forward Reference)[5]라하며, 최대 순방향 참조를 MF라고 부른다.

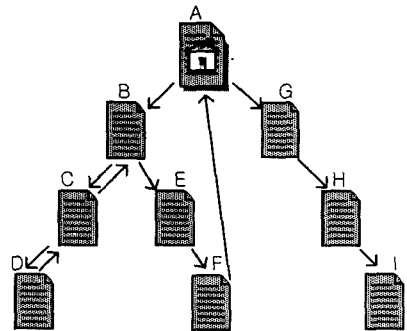


그림 2. 순회패턴(Traversal Pattern)의 예

예를 들어 사용자가 그림 2와 같은 {A, B, C, D, D, C, B, E, F, A, G, H, I} 순서로 웹 사이트를 방문한다고 가정하자. 이런 경우, 순회패턴 알고리즘을 적용하면, {ABCD, ABEC, AGHI}와 같은 최대 순방향 참조를 얻을 수 있다. 즉, 사용자가 웹을 검색함에 있어서 역방향 링크를 다시 검색할 경우 순방향 참조가 종료되고 하나의 최대 순방향 참조를 만들게 된다. 알고리즘을 통해 순방향 참조 집합을 구할 때 각 노드에 해당하는 링크 앵커 문자 키워드를 추출한다. 링크 앵커는 <a>태그와 </a>태그 사이에 존재하는 문자로, 예를 들면 <a href="adtl.ajou.ac.kr">Ajou university</a>같은 태그가 존재할 경우, 'Ajou university'를 앵커 문자 키워드라 한다. 링크의 앵커 텍스트는 링크에 대한 가장 중요한 정보들을 포함하고 있다. 이런 방식으로 추출한 키워드와 웹 사전인 Open Directory Project(ODP)[6]으로부터 얻은 16개의 카테고리 고리사이의 유사도를 측정해서, 사용자들의 관심 영역을 추출할 수 있다. 사용자들이 주로 접근하는 MF와 관심영역을 고려해서 의미있는 MF인 Sitemarking links를 4.2절의 모델을 통해서 얻을 수 있다.

#### 4.2 Sitemarking links 모델

제안하는 기법은 사용자의 MF접근빈도와 관심영역을 바탕으로 의미있는 MF인 Sitemarking links를 추출하게 된다. 이것의 추출을 위해 사용되는 가중치 계산을 위한 Sitemarking links모델은 다음과 같다.

사용자들의 집합  $U = \{u_1, u_2, \dots, u_n\}$ 을 n명의 사용자들의 집합이라 하자. 각각의 사용자  $u_n$ 은 사이트 단위의 북마크의 링크들(L)과 흥미 카테고리(C)로 구성되어 있다. 사이트 단위의 링크들인 L은 자주 접근한 순서로 정렬되어 있으며, 관심 카

태고리는 관심영역 추출기법에 의해 추출된다. 가중치 계산을 위한 변수들을 정리하면, 다음 표1과 같다.

표 1. *Sitemarking links*을 구하기 위한 변수들

$U$ 는 $n$ 명의 사용자들의 집합
모바일 이용자 $u$ 는 자주 접근한 링크 ( $L$ )과 흥미있는 카테고리 ( $C$ )로 구성
링크 집합 $L$ 은 많이 접근한 횟수에 의해 정렬되어 있음

식(1)과 같이 *Sitemarking links*은 북마크 링크와 관심영역의 유사도로 구성되어져있다. 여기서  $\tau$ 는 임의의 기준값이다. 관심 영역의 유사도는 식(2)와 같이 사용자들을 조사하면서 최대의 유사도값을 추출한다.

$$S = \{ \langle l, weight(c) \rangle | l \in L, weight(c) \geq \tau \} \quad (1)$$

$$weight(c) = \sum_{i=1}^n Max(C) \quad (2)$$

5. 사용자 모델링

현재 무선단말기 종류는 다양하며, 통일된 표준은 없다. 그렇기 때문에 본 논문에서는 모바일 기기들의 특성을 고려해서 정보를 제공하기 위해 CC/PP[7]을 이용한 정적 프로파일을 생성한다.

5.1 정적 프로파일(Static profile)

모바일 사용자가 프록시인 에이전트 서버(agent server)에 처음 접속하면, 사용자의 모바일 기기의 정보를 보낸다. 이 설명도는 XML형태인 CC/PP(Composite Capabilities/ Preference Profiles)[7]을 이용해서 보낸다. CC/PP는 현재 W3C에서 표준화가 진행 중이며, 서로다른 모바일 디바이스에 알맞은 정보를 제공하기 위해서 기기의 하드웨어, 소프트웨어 그리고 사용자의 선호도 등을 기술해 준다. 그림 3은 제안하는 논문에서 이용하는 CC/PP의 예이다. 크게 기기, CPU, 메모리, 자바가상머신의 버전, 모바일코드를 전송받는 방식인 EMAIL 그리고 사용자의 관심분야의 총6가지로 구성되어있다. 디바이스 부분은 기기의 이름과, 운영체제버전, 그리고 화면의 사이즈로 구성되어 있으며, CPU와 메모리는 하드웨어 용량을 묘사하고 있다. 또한 플랫폼에 상관없이 동작하기 위해서 자바가상머신의 버전을 명시했다. 에이전트 클라이언트가 다운받아지는 방식에 대해서 묘사했다. 그림 3에서는 이메일을 통해서 받는다. 마지막으로 사용자의 관심분야를 표시하는 부분이 있다.

5.2 사이트 단위의 북마크 추천기법

모바일 사용자가 프록시에 *Sitemarking links*에 없는 정보를 요청시, 모바일 에이전트는 프록시인 모바일 서버에 접속해서 필요한 정보를 가져온다. 하지만 프록시에 요청한 사이트에 대한 정보가 있다고 무조건적으로 가져올 순 없기 때문에, 유사도를 측정해서 관심분야의 사용자의 정보를 가져온다. 모바일 에이전트는 여러 에이전트 서버를 돌아다닐 수 있다. 이렇게 돌아다니면서 작성한 CC/PP의 Preference항목의 값을 추출해서 유사도를 측정한다. 측정된 유사도의 값이 전체 값의 반보다 높은경우, 추출한 다음 요청한 사용자에게 제공한다.

```
<?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">
.....
<ClientCapabilities>
<Device>
<Name></Name>
<OS>Window CE</OS>
<ScreenSize>320*240</ScreenSize>
</Device>
<CPU>
<Speed>300MHz</Speed>
</CPU>
<Memory>
<RAM>16MB</RAM>
</Memory>
<JavaVirtualMachine>
<Version>1.0</Version>
</JavaVirtualMachine>
<Download>
<Type>Email</Type>
<Info>seaaes3@ajou.ac.kr</Info>
</Download>
<Preference>
<Preferences>Computer</Preferences>
</Preference>
</ClientCapabilities>
</rdf:RDF>
```

그림 3. 모바일 사용자를 위한 CC/PP의 예

6. 결론

무선 단말기 사용자의 인터넷 접속이 증가하면서 웹 페이지는 유선 사용자뿐만 아니라 무선 사용자의 요구를 만족시켜야 할 필요성이 증가했다. 이러한 필요성으로 인해 현존하는 웹 페이지들이 각각의 기기에 맞게 표현되어야 하지만, 무선 단말기는 데스크톱에 비해 제한된 인터페이스와 화면 크기를 가지므로 사용자에게 사용의 불편함을 가져온다.

본 논문에서는 기존의 페이지 단위의 북마크 방식을 개선한 사이트 단위의 북마크를 제공해줌으로써, 모바일 장치 사용자들의 탐색편의를 높여준다. 또한 모바일 에이전트의 개념을 이용해서 관심분야가 유사한 사용자들이 해놓은 사이트 단위의 북마크를 추천해준다.

7. 참고문헌

- [1] <http://www.comscore.com/> -- comScore Networks, Inc.
- [2] Timothy W. Bickmore and etc., " Digestor: Device-independent access to the World Wide Web". In Proceedings of the 5th WWW Conference, 1996
- [3] Orkut Buyukkokenm and etc., " Power Browser: Efficient Web Browsing for PDAs", Digital Libraries Lab, Stanford University, 2000
- [4] Debbie O. Chyi, " An Infrastructure for a Mobile-Agent System that provides personalized services to Mobile device," Technical Report TR2000-370, Dartmouth College, Computer Science, Hanover, NH, 2000
- [5] Ming-Syan Chen, and etc., " Efficient Data Mining for Path Traversal Patterns," IEEE Transactions on Knowledge and Data Engineering, Vol 10, No. 2
- [6] <http://dmoz.org> - Web directory of Open Directory Project
- [7] <http://www.w3.org/Mobile/CCPP/>