

북 마크를 이용한 사용자 행태 분석 시스템의 설계 및 구현

권재호, 김계숙, 박규석
경남대학교 컴퓨터공학과

Design and Implementaioin of a User's Behavior System using Bookmark

Jae-Ho Kwon, , Gye-Sook Kim, Kyoo-seok Park
Dept. of Computer Science, Kyungnam University
E-mail : jhavarado@korea.com, gskim@dongju.ac.kr, kspark@kyungnam.ac.kr

요 약

인터넷을 항해하는 동안, 사용자들은 관심있고 흥미있는 사이트들을 자신의 즐겨찾기 폴더에 저장하여 다음에 쉽게 접속을 할 수 있도록 한다. 왜냐하면, 수많은 사이트들을 기억할 수 없기 때문이다. 본 논문에서는 북마크 기능을 이용하여 단순히 즐겨찾기 폴더에 저장된 사이트들을 보여주고 관리하는 사용자관점이 아닌, 서버측면에서 사용자의 행태를 분석하는 방법을 제안하고자 한다.

1. 서론

인터넷에는 수많은 정보들이 저장되어 있다. 인터넷의 발전과 더불어 갈수록 증가되는 정보들로 인해 인터넷 사용자들은 북마크 기능을 이용하여 자신만의 즐겨찾기 폴더를 구성할 필요성을 느끼게 되었다. 수많은 사이트들이 있고, 자신이 방문한 사이트라 할지라도 다음에 재방문을 하려면, 사이트의 이름을 기억하지 않는 이상, 검색엔진으로 다시 찾아야 되는 불편함이 있기 때문이다. 이러한 경향은 개발자들로 하여금, 사용자들이 북마크 기능을 이용하여 저장해 놓은 사이트 및 웹 문서들을 보여주고 관리할 수 있는 기능을 제공해 주고, 사용자들이 저장해 놓은 웹 사이트들의 존재유무를 체크해 주는 기능을 가진 툴을 제공해왔다.

기존의 북마크 기능을 이용한 방법은 사용자관점에서 주로 즐겨찾기 폴더에 저장된 사이트나 웹 문서들을 보여주고, 윈도우 탐색기처럼 폴더를 생성, 삭제하는 등의 관리하는 기능을 제공하며, 저장된 사이트나 웹 문서들의 존재유무를 체크해 주는 기능을 제공한다.

본 논문에서의 분석시스템은 사용자관점이 아닌, 서버측면의 사용자 행태분석을 중심골격으로 하고 있다.

인터넷을 항해하다가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 사이트 및 웹 문서들을 이용하여 사용자의 행태를 분석한다. 제안 시스템은 방문한 사이트와 웹 문서들을 주기별 통계(등록 사이트 증감 현황, 즐겨찾는 사이트 통계, 웹 사이트별 웹서버 통계), 일시적 통계(등록 사이트 수, 카테고리 수, 데드 링크 현황)등을 통해 서버측면에서의 사용자 행태를 분석하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 북마크 기능에 관련된 연구에 대해 살펴보고, 3장에서는 본 논문에서 제안하는 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 사이트 및 웹 문서들을 통해 사용자의 행태분석을 위한 시스템 설계를 하고, 4장에서는 구현 및 평가, 5장에는 결론을 내린다.

2. 관련연구

2.1 웹 자원의 구성

사이트나 웹 문서들의 위치, 방문 횟수, 최근 사용한 날짜, 개별적인 사용자의 선호 등의 다양한 요소들은 문서들을 분류 및 구성하고 사용자의 행태를 알아보는 중요한 요소들이 된다. 하나의 요소만으로는 충분하지 않으며, 효과적인 사용자의 정보를 접근하기

위해서는 일반적으로 여러 가지 요소들이 고려되어야 한다. 따라서, 효과적으로 사용자의 정보를 관리하는 틀을 제공하려면, 인덱싱(Indexing), 정렬 등의 기능을 지원하고 결합해야 한다.

2.2 저장과 인덱싱

북마크의 저장과 인덱싱 메커니즘은 각각의 사이트나 웹 페이지가 격장모양의 구조로 카테고리의 결합과 함께 표현된다[1].

2.3 북마크 불러오기

인터넷 익스플로러의 북마크 구조를 불러오기 위해, 북마크 구조를 파싱하고, 그와 동일한 구조의 북마크를 생성한다. 시스템은 초기의 카테고리 집합을 생성하기 위해 계층적인 즐겨찾기 폴더의 이름들을 사용한다[1].

2.3 사용자 북마크의 이용

사용자의 즐겨찾기 폴더는 웹 페이지에 대한 사용자의 행태를 알 수 있는 중요한 단서가 된다. 사용자의 즐겨찾기 폴더 조직을 보면 관련 사이트 및 웹 페이지의 정보를 얻을 수 있다. 그림 1에 사용자가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 일반적인 형태를 나타내었다.

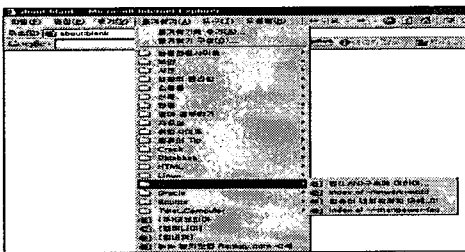


그림 1. 사용자의 즐겨찾기 폴더 구성도

북마크는 최상위 분류 아래 1차 폴더, 2차 폴더, n차 폴더를 가질 수 있다. 사용자는 자신의 정보요구에 따라 분류항목을 생성하고 그룹화 한다. 인터넷에 연결된 사용자는 웹 브라우저를 통해 웹 페이지를 보고 자신이 필요로 하는 정보가 포함된 웹 페이지는 북마크에 추가하고, 기존의 웹 페이지에 대하여 사용자의 취향이 변화하거나 정보 가치가 하락하면 해당 페이지를 삭제한다[2].

3. 시스템 설계

북마크 기능을 이용하여 개인의 즐겨찾기 폴더의 링크존재 유무를 확인할 수 있고, 사용자가 북마크 기능을 사용한 시간 등을 보여줌으로써, 서버측면에서의 기능을 강조하였다. 또한, 최근 방문한 사이트를 몇 개 단위로 보여주기, 가장 많이 방문한 사이트로 정렬하여 보여주기, 해당 사이트나 웹 페이지가 있는 OS와 웹 서버, Browser 정보를 보여줌으로써, 다양한 정보분석을 제공하여 사용자 행태에 대해 북마크 기능을 이용하여 분석할 수 있다.

제안 시스템은 사용자가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 내용을 보여주는 뷰(View) 모듈, 북마크 정보를 데이터베이스에 저장하는 모듈, 북마크를 사용함으로써 증가되는 사이트들의 증감현황과 즐겨찾는 사이트의 통계(Top 10/50), 각 사이트별 방문 횟수 및 방문 페이지 수 및 이용시간에 대한 통계를 보여주는 분석 및 통계모듈, 웹 사이트별 사용하는 웹 서버와 링크의 유무존재, 주간 통계, 월간 통계, 연간 통계를 관리하는 스케줄 관리모듈로 구성되어 있다. 제안 시스템의 전체 구성도는 그림2와 같다.

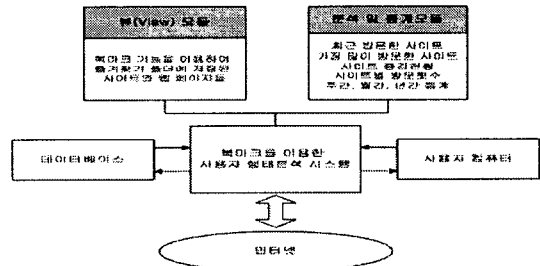


그림 2. 사용자 행태분석 시스템의 구성도

3.1 뷰(View) 모듈

뷰(View) 모듈은 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 내용을 보여주는 모듈이다. 사용자가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 사이트나 웹 페이지가 데이터베이스에 구성된 통계 및 분석모듈과 함께 뷰(View) 모듈과 연관되어 나타난다.

3.2 분석 및 통계모듈

분석 및 통계모듈은 사이트나 웹 페이지를 방문한 횟수를 주기별로 나타내고, 등록 사이트의 증감현황 등을 나타낸다.

◆ 주기별 통계

- (1) 등록 사이트 증감 현황
 - 등록사이트 증감현황 : 주간, 월간, 연간
 - 신규등록 사이트 수 : 지난번, 지난달, 지난해
 - 삭제사이트 수 : 지난번, 지난달, 지난해
- (2) 즐겨찾는 사이트 통계(10/50개까지) 각 사이트 별 방문횟수 및 방문 페이지 수 및 이용시간
 - 주간 통계, 월간 통계, 연간 통계
- (3) 웹 사이트별 사용 웹 서버 통계
 - 분석 날짜별
 - 월별, 년별

◆ 일시 통계

- 등록 사이트 수
- 카테고리 수 : 1차, 2차, 3차 등...
- 링크존재유무 현황
- 웹 서버의 종류
- 웹 브라우저의 종류

분석과 통계모듈은 북마크를 이용하여 즐겨찾기 폴더에 저장한 사이트나 웹 페이지를 기준으로 방문한 시간과 횟수, 사용자가 매일 사이트를 추가한 정보, 삭제한 횟수와 시간, 그 사이트의 링크존재유무, 웹 서버 종류등의 정보를 분석하여 뷰(View) 모듈에 보여주기 위해 필요한 정보들이다. 전체적인 구성도를 보면, 그림3과 같다.

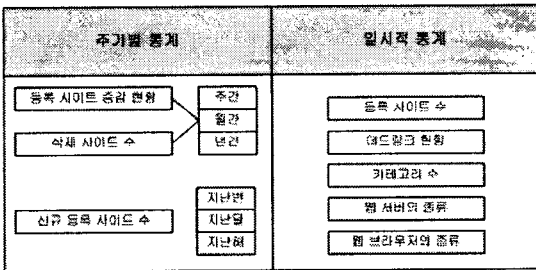


그림 3. 분석 및 통계모듈의 구성도

3.4 데이터베이스 설계

북마크 기능을 이용한 사용자행태 분석을 위한 시스템은 즐겨찾기 폴더의 내용과 방문한 시간, 웹 서버의 종류, 브라우저의 종류등의 정보를 데이터베이스에 저장하고 뷰(View)모듈에서 그 정보를 보여준다. 사용자행태를 분석한 정보를 보여주기 위해 각 카테고리마다 데이터베이스를 구축하여 서버측에서의 통합

관리를 통하여 정보를 확인할 수 있도록 한다. 사용자가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 사이트나 웹 페이지들을 새로 재구성하여 결과를 보임으로써, 북마크에 대한 흐름을 확인할 수 있다. 전체 데이터베이스 설계의 구성도는 그림4과 같다.

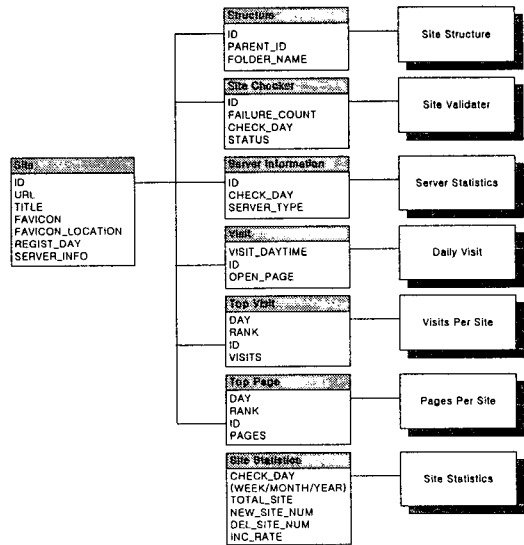


그림 4. 데이터베이스 구성도

4. 구현 및 평가

본 연구에서는 사용자가 북마크 기능을 이용하여 즐겨찾기 폴더에 저장한 사이트 및 웹 페이지를 분석하는 분석 및 통계모듈과 뷰(View)모듈, 데이터베이스를 구성하는 모듈로 구성된다.

제안된 시스템에 대한 운용환경은 표 1과 같다.

표 1. 시스템 하드웨어 및 소프트웨어 환경

분류	명세
Server System	
Server OS	Windows 2000 Server
Web Browser	Internet Explorer 6.0
DBMS	Microsoft Access 2002
개발 Tools	Visual Basic 6.0

즐거찾기 폴더가 적용된 메인 화면은 그림 5와 같다.

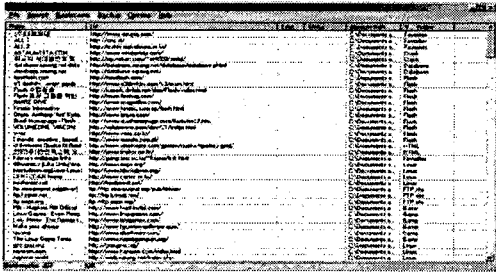


그림 5. 즐거찾기 폴더의 메인 화면

5. 결론

본 논문에서는 기존의 사용자관점의 북마크 툴을 특정시간 및 정보를 추가하여 사용자가 북마크 기능을 이용하여 즐거찾기에 저장한 사이트나 웹 페이지를 분석하여 DB를 구축하고 서버측면에서 정보들을 볼 수 있는 행태분석 시스템을 제안하였다. 그러나, 여러 사용자가 URL을 공유하거나 사이트 및 웹 페이지를 추천할 수 있는 알고리즘 등은 계속하여 연구할 예정이다. 제안시스템은 검색엔진분야에 적용될 수 있으리라 생각하며, 특히 북마크를 이용해서 사용자 취향 및 성향을 분석하는데 일조를 하리라 생각한다.

[참고문헌]

- [1] Richard M.keller, "A bookmarking service for organizing and sharing URLs", Computer Networks and ISDN System 29(1997) 1103-1114
- [2] 安柄鮮, "An Intelligent Web Agent using Bookmark Monitoring and Multi-Profile", 東國大學校 碩士學位, 1997. 12
- [3] Peter Edwards, David Bayer, Claire L. Green, Terry R. Payne, "Experience with Learning Agents which Manage Internet-Based Information", AAAI Spring Symposium on Machine Learning in Information Access, 1996.
- [4] James Ruker, Marcos J. Polanco, "Sitseer: Personalized Navigation for the Web", Communications of the ACM, pp. 73-75, March 1997/Vol. 40, No. 3.
- [5] A. Jennings, H. Higuachi, "A personal news service based on a user model neural network", IEICE Transactions on Information Systems, 75(2):198-209, March 1992.

[6] G. Fischer, C. Stevens, "Information Access in Complex, Poorly Structured Information Spaces", In Human Factors in Computing Systems, CHI '91 Conference Proceedings, ACM, 1991.

[7] B. Sheth, P. Maes, "Evolving Agents for Personalized Information Filtering", Proceedings of the Ninth Conference on Artificial Intelligence for Applications, pp.345-352, IEEE Computer Society, 1993.