폐쇄형 브라우저 시스템 설계 및 구현

임 범 춘*, 황 명 국*, 이 경 현**
부경대학교 *전산정보학과, **전자계산학과

Design and Implementation of the Private Browser System

Bum-chun Lim*, Myung-gook Hwang*, Kyung-hyun Rhee**
*Dept. of Computer Information, **Dept. of Computer Science PuKyong Nat'l University.

요 약

우리는 정보를 찾기 위해서 신문, 도서관보다는 인터넷을 먼저 검색한다. 이와 같은 사회환경에서 불건전 정보 증가는 이전 사회에서 인식하지 못했던 문제를 야기하고 있다. 그 대상이 청소년뿐만 아니라 아동까지 확대되고 있어, 문제는 더운 더 커질 수 있다. 이러한 문제점을 극복하기 위해 본 논문에서는 연령별 또는 등급별 웹브라우저를 통하여 접근할 수 있는 범위를 제한하여 전문 검색 에이전트를 통하여 검색과 카테고리 분류를 하여 DB구축을 제안하고, 네트워크의 데이터링크 계층에 속하는 NIC(Network Interface Card)의 드라이버 수준에서 건전 사이트만을 접근이 가능하도록 폐쇄형 브라우저 시스템의 설계 및 구현방법에 대해서 논의한다.

1. 서론

현재 인터넷은 하나의 사회현상에서 문화 및 생활의 일부로 자리잡고 있으며, 인터넷의 확장은 현실에서 발생하는 여러 가지 역기능 또한 온라인 공간으로 고스란히 옮겨 놓고 있다. 유해정보에 대한 사회적 우려와 요구는 국내·외적으로 문제로 논란이 되고 있다. 예를 들어 베텔스만 재단이 1999년 6월에 조사한 결과에 따르면, 독일인의 81%(90%), 미국인의 85%(86%), 호주인의 76%(78%)가 인터넷 이용과 관련된 위험을 우려하고 있는 것으로 조사되었고(괄호의 안의 수치는 인터넷 이용 경험이 있는 응답자임)[1], 우리나라에서는 최근 서울지역 초등생 1천135명을 대상으로 설문조사를 실시한 결과 '음란, 엽기 등 불건전 사이트 접속 경험'을 물음에 대해 '가끔 들어갔다', '자주 들어갔다', '항상 들어갔다'는 응답이 82.6%가 '인터넷을 하다가 우연히 들어가게 됐다'고 답했고, 사이트의 이름을 알고 찾아 들어갔다'는 초등생도 17.4%로 의외로 많았다[2]. 이와 같이 유해정보에 초등학생들에게까지 노출이 되어 있는 상태이다. 그리고, 인터넷은 전 세계적으로 연결되어 사용되는 개방망의 특성을 가짐에 따라 유해 정보 제공자를 각국의 법적, 제도적 장치를 이용하여 규제하는데는 한계가 있으므로, 인터넷상의 유해 정보로부터 청소년들을 효과적으로 보호하기 위한 기술적인 대책이 시급히 마련되어야 한다. 기술적인 대책으로 정보통신 선진국에서는 접속차단 소프트웨어를 개발하여 해결하고 있다.

본 논문에서는 응용계층의 소프트웨어 기반이 아닌 하드웨어 방식의 시스템 방식을 제안하려고 한다. 이 방식은 전문 검색 에이전트를 통하여 도메인을 수집하고, 분류 에이전트를 이용하여 유해목록과 허용목록을 구분한 뒤, DB관리 서버에 저장을 한다. DB관리 서버에 분류된 목록들 중에 허용목록으로 판별된 사이트들만 접근 모듈을 통하여 접근이 가능하도록 한다. 접근 모듈은 드라이버 수준에서 구현되어 사용자에게 투명한 환경을 제공하게 되며 목록을 별도로 관리하지 않아도 된다.

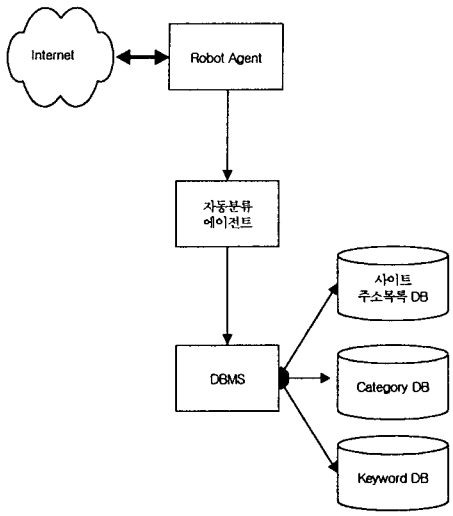
본 논문의 구성은 먼저 2장에서 제안하는 에이전트들에 대한 기술 설계를 설명하며, 3장에서는 제안하는 폐쇄형 브라우저 시스템에 대한 기술 설계를 논의하여 기존 방법과 차별되는 시스템 구현 부분과 유효성에 대하여 기술한다. 마지막 4장에서는 결론과 향후 연구과제를 논의한다.

2. 전문 검색 에이전트와 분류 에이전트에 대한 기술 설계

1) 전문 검색 에이전트의 대한 기술 설계

전문 검색 에이전트란 인터넷이라는 광대한 영역을 수집 목적에 따라 검색하고, 자동적으로 URL을 추출하는 방식으로 작동하는 에이전트이다. 검색 기준은 카테고리 분류에 대표사이트인 야후(http://www.yahoo.com)에서 처음 행하였던 것으로 교과과정별로 URL을 분류해 각 목록들을 만들어 나갈 것이다[3]. 분류된 목록을 기준으로 전문 검색 에이전트를 통해 URL 목록들을 자

동 분류 에이전트에게 전달하여 DB에 카테고리 별로 저장한다. 이 것을 토대로 모듈 매니저에 의해서 허용목록 범위를 결정하여 결정된 범위 내의 사이트만 접근하는 방법을 제안하였다.



(그림 1) 전문 검색 에이전트 구성도

(그림 1)은 전문 검색 에이전트에서 사용되는 구성요소들로 기능들을 다음과 같다.

- Robot Agent : 인터넷 상에서 관련 수집 정보들을 수집해 오는 역할을 담당 에이전트를 말한다.
- 자동 분류 에이전트 : 분류 기준에 맞게 분류하는 기능을 가진 에이전트로 DBMS(Database Management System)에게 수집된 정보를 분류해서 넘겨주는 역할을 담당하는 에이전트를 말한다.
- DBMS(Database Management System) : 사이트 주소 목록 DB, Category DB, Keyword DB등을 관리하는 System을 말한다.
- 사이트 주소 목록 DB : 유해한 것과 허용할 data를 구분하여 저장하는 저장소를 말한다.
- Category DB : 사이트 목록별 카테고리를 분류해 data를 저장하는 저장소를 말한다.
- Keyword DB : 수집된 keyword들에 정보를 저장하는 저장소를 말한다.

2) 분류 에이전트에 대한 기술 설계

분류 에이전트는 카테고리의 Keyword 및 가중치 정보와 에이전트에서 가져온 문서에서 추출한 단어들을 이용하여 문서의 카테고리에 대한 유사도를 계산하여 허용할 사이트와 유해 사이트로 구분되어 DB(Database)에 저장이 된다.

분류 에이전트에서 사용되는 구성요소들은 아래와 같다.

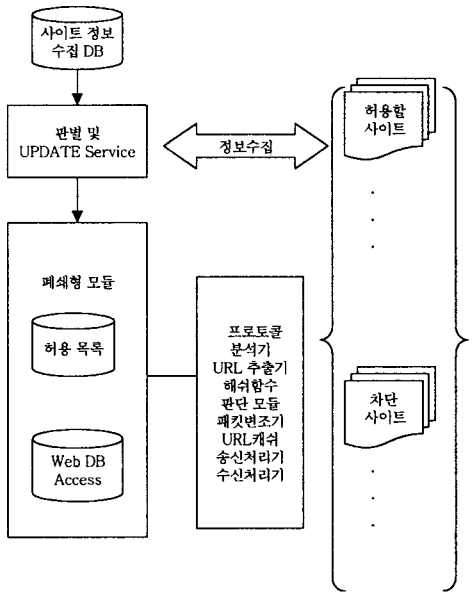
- 웹 문서 : Robot Agent를 통해 수집해 온 문서들을 말한다.
- 형태소 분석 : 웹 문서에 제공되는 문구들의 형태소를 분석하는 것을 말한다.
- 가중치 부여 : 형태소 분석을 통해 구분된 문구들의 일정 기준값을 부여하는 것을 말한다.
- 유사도 계산 : 부여된 가중치 정보를 계산한 결과값이 판단의 기준이다.

3. 폐쇄형 브라우저 시스템 설계 및 구현

본 장에서는 2장에서 설계된 에이전트를 이용하여 유해 목록과 허용 목록을 구분하도록 설계되었다.

위 설계를 토대로 클라이언트에서는 허용 목록만을 접속하도록 구현하고, 어플리케이션이 아닌 드라이버 수준에서 구현한 보안시스템을 설계하였다.

1) 폐쇄형 브라우저 시스템의 구조



(그림 2) 전체 시스템 구조

(그림 2)는 전체 시스템 구조로 본 논문에서 구현한 모듈은 Microsoft사의 Windows 운영체제를 기반으로 하며 ANSI C로 구현하였다. 드라이버는 Microsoft사의 Windows 운영체제상에서 동작되는 커널 모드 드라이버인 miniport 드라이버이다[4]. miniport 드라이버의 특징은 하드웨어를 드라이버 제작자가 직접 제어할 수 있다는 것이다. 따라서, NIC와 직접적으로 연결되어 NIC를 제어할 수 있으며 표준 하드웨어 이외에도 개발자가 정의하는 하드웨어를 사용할 수 있다. 본 논문에서는 별도의 플래쉬 메모리를 사이트의 목록을 저장하는 용도로 NIC에 장착하였다.

폐쇄형 브라우저 시스템은 차단 모듈에서는 허용목록을 가지고 있으며 모듈 매니저에서는 차단 목록을 DB로 구축하여 사용한다.

본 시스템에서 사용되는 구성요소들의 기능은 아래와 같다.

① 허용 모듈

허용 모듈의 동작은 드라이버가 송·수신 패킷의 처리를 마치기 직전에 패킷을 허용 모듈로 보내어 허용모듈에서 패킷을 처리하는 방법을 이용한다.

- 허용 목록 : NIC에 장착된 플래쉬 메모리에 저장되는 목록이다. URL이 저장되어야 하지만 길이가 URL마다 다르므로 해쉬함수를 이용하여 8byte의 고정된 길이로 만들어 저장하였다.
- WEB DB Access : 질의를 할 수 있는 모듈 매니저 서버의 정보를 가지고 있다.
- URL 추출기 : HTTP 헤더에서 URL부분을 추출하는 모듈이다[9].
- 해쉬함수 : 추출된 URL을 8바이트의 해쉬 값으로 만든다. 사용 알고리즘은 MD5이다[11].
- 프로토콜 분석기 : 송·수신되는 패킷들의 프로토콜을 분석한다[5,9].
- 판단 모듈 : 사용자가 접근하려는 사이트의 접근 허용 여부를 판단한다. 판단은 추출된 URL의 해쉬를 취한 값이 플래쉬 메모리에 기록되어 있는 허용목록에 있는지 검색하여 판단한다[5,8].
- 패킷 변조기 : 판단 모듈에서 모듈 매니저 서버로 질의를 요구한다면 HTTP request 패킷을 모듈 매니저 서버로 전송되도록 변조하여 질의에 필요한 데이터를 패킷에 삽입한다.
- 송신 처리기 : 드라이버의 송신부로 패킷을 내보낸다[4].
- 수신 처리기 : 드라이버의 수신부로 패킷을 내보낸다[4].

2) 모듈 매니저서버

서버측의 구성은 폐쇄형 모듈을 관리하고 허용 목록과 유해 목록의 지속적인 업데이트가 가능하도록 설계한다.

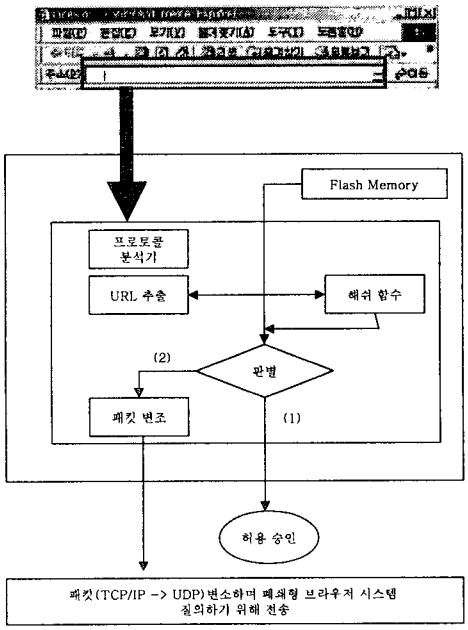
- 모듈 매니저 : 폐쇄형 모듈에서의 질의를 처리하거나 폐쇄형 모듈의 허용목록 업데이트와 서버측 유해목록의 업데이트를 수행한다. 유해목록의 수집은 검색 에이전트를 이용하여 수집하다. 모듈 매니저의 역할은 폐쇄형 모듈을 관리하는 정책을 결정하는 것이다. 정책의 내용은 폐쇄형 모듈의 허용목록 업데이트 또는 제거 방법, 유해목록의 업데이트 또는 제거 방법, 폐쇄형 모듈에 적용할 카테고리의 범위 결정 등의 정책을 수립하여 모듈 매니저에 적용하여야 한다.

3) 폐쇄형 브라우저 시스템의 구현

폐쇄형 브라우저 시스템은 상술한 miniport 드라이버의 모듈로 구현되며 사용자의 HTTP request를 감지하면 URL을 추출하여 플래쉬 메모리의 URL들과 비교를 하게 된다. 본 논문에서는 사용되는 방법은 허용목록에 있는 사이트만을 통과하도록 되어 있으며, 모듈관리자에 의해 허용목록이 업데이트 되도록 설계되었다. 송·수신되는 패킷들은 허용 모듈에서 감지해야 하는 HTTP패킷이거나, 모듈 매니저의 응답이거나 모듈에서 처리가 필요하지 않는 패킷일수도 있다. 그러므로, 패킷의 프로토콜을 분석하여 적절한 처리를 하여야 한다. (그림 3)은 송신과정에서의 허용모듈의 동작과정을 설명하고 있다. (그림 3)에서 나타나듯이 사용자의 HTTP request 패킷을 허용모듈에서 인식하여 사용자가 접근하고자 하는 목적지 사이트의 URL을 추출하여야한다.

본 논문에서 구현한 URL 추출기는 현재 널리 사용되고 있는 웹 프락시(Proxy) 서버를 통한 트래픽 제어가 가능하도록 설계가 되었다. 웹 프락시를 통한 트래픽은 HTTP request부터 시작하는 것이 아니라 웹 프락시에 목적지 사이트의 URL을 질의어나 데이터로 전송하는 방식을 취한다.

① 폐쇄형 모듈의 송신시 동작

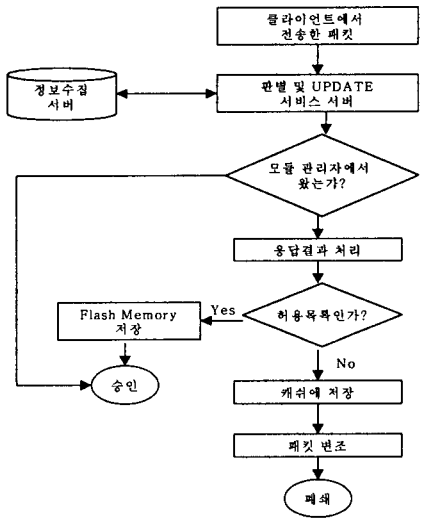


(그림 3) 송신시 동작과정

(그림 3)의 (2)의 동작은 폐쇄형 모듈의 판단 모듈에서 판별결과가 모듈 매니저 서버로의 질의를 요구하고 있으며 사용자의 HTTP request 패킷을 UDP(User Datagram Protocol) 패킷으로 변조하여 판단에 필요한 데이터를 삽입하여 서버로 보내게 된다. UDP를 사용하면 패킷이 유실 될 수 있으나 HTTP request 패킷을 보낸다는 것은 사용자의 PC와 사용자가 접근하려는 웹 서버사이에 TCP의 three-way hand shake가 성공적으로 수행되었다는 의미이기도 하다.

따라서, 운영체제의 TCP(Transmission Control Protocol)/IP(Internet Protocol) 스택에서는 드라이버에 정상적인 HTTP request 패킷이 전달되는 것으로 여기게 된다. 그러므로, 모듈 매니저 서버에 보낸 UDP 패킷이 유실되어도 운영체제의 TCP/IP 스택의 재전송 메커니즘을 이용할 수 있게 된다. 그리고, UDP의 속도는 TCP에 비해 빠르다[10]. (그림 3)의 (1)의 동작은 판단 모듈의 판별 결과가 허용목록 판단된 결과이며 HTTP request 패킷을 정상적으로 처리하도록 한다.

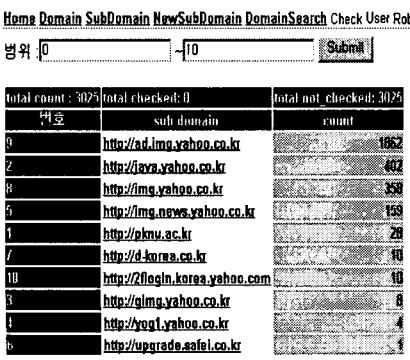
② 차단 모듈의 수신시 동작



(그림 4) 허용모듈의 수신시 동작과정

(그림 4)는 허용 모듈의 수신시 동작을 설명한다. 수신시 처리가 필요한 부분은 모듈 매니저에서 응답 패킷만을 처리한다. 본 논문에서 제안한 폐쇄형 브라우저 시스템의 허용 모듈은 HTTP request 패킷은 유해사이트에 도달하지 못한다. 따라서 HTTP request 패킷을 변조하여 모듈 매니저로의 질의를 수행하게 되며 질의에 대한 모듈 매니저의 응답 UDP 패킷을 수신과정에서 처리한다. 즉 수신되는 HTTP 패킷이나 다른 용도의 패킷을 받았을 경우 (그림 4)와 같은 처리를 하게 되며 응답 패킷의 판단결과 데이터가 허용목록인 경우는 원하는 사이트가 open 될 것이고, 유해목록일 경우 응답 UDP패킷을 사용자가 접근하려던 웹 사이트의 응답 HTTP redirection 패킷으로 변조하여 TCP/IP 스택에게 세션의 응답으로 보내주고 임시 메모리에 URL의 해쉬값을 기록한다. 응답 패킷의 판단결과 데이터가 허용사이트인 경우 해당 사이트의 URL의 해쉬값을 플래쉬 메모리에 기록하게 된다.

이와 같은 결과로 (그림 5)에서는 폐쇄형 브라우저 시스템의 DB관리 모듈에서의 목록들을 보여 주고 있다.



4. 결과 및 향후 연구과제

지구상의 환경오염문제와 함께 대두되고 있는 것이 인터넷상의 환경 또한 사회문제로 야기되고 있다. 폐쇄형 브라우저 시스템은 기존 차단 시스템의 역기능을 구현하여, 다음과 같은 특징과 보안점이 필요하다. 첫째, 폐쇄형 브라우저 시스템은 관리가 용이성이다. 본 논문에서 구현한 시스템은 허용목록이 지속적으로 업데이트 되도록 설계되어 있으며 목록의 작성도 실시간으로 사용자가 자주 방문하는 사이트를 위주로 작성하도록 구현되어 있다. 둘째, 사용자의 시스템에서 투명하게 동작한다. 셋째, 에이전트들의 인공지능화와 자연어 분석기, 자동 학습 기능 등을 도입은 현 서버 시스템의 기능적, 질적인 향상을 가져 올 수 있다. 넷째, 호환 드라이버가 아닌 전용 NIC를 가지는 드라이버 방식으로 제작한다면 성능의 향상을 가져올 수 있다. 기존 많은 차단 프로그램과는 비교는 배제하였다. 네트워크와 PC의 성능 향상에 따라 본 논문에서 제안한 폐쇄형 브라우저 시스템은 구현된 모듈관리자는 DB의 관리와 갱신, 그리고 정보의 수집에 대한 보다 나은 방법이 많이 요구되고 있다. 또한 관리정책에 따라 사용자에게 보다 질 좋은 서비스가 가능하므로 이에 대한 추후 보완 방안도 보안되어야 할 것이다.

[참고문헌]

[1] 정보통신 정책 자료집 中, "인터넷 컨텐츠 자율규제의 개념과 구성", 인터넷자율규제포럼 R3net, p4 2000.

[2] 오연주, "초등학교 학생의 정보통신윤리 의식에 관한 조사연구", 서울교육대학원 p39-46 2002.

[3] 오문길 역저, "야후! 성공방식" 물푸레출판사, p10-20, 2001.

[4] Microsoft coporation, "Driver Developent Toolkit", http://msdn.microsoft.com, June 2000.

[5] Postel J.,"Internet Protocol", RFC 791, USC/Information Sciences Instutute, September 1981.

[6] Reynolds, J., Postel, J., "Assigned Numbers", RFC 1340, USC/Information Sciences Instutute, July 1992.

[7] Postel, J., "User Datagram Protocol", RFC 768, USC/Information Sciences Instutute, August 1980.

[8] Postel, J., "Transmission Control Protocol", RFC 761, USC/Information Sciences Instutute, January 1980.

[9] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Berners-Lee, T., "Hypertext Transfer Protocol --HTTP/1.1", RFC 2068, January 1997.

[10] Richard Stevens, W., "TCP/IP Illustrated: the protocol", Addison-Wesely, 1994.

[11] Rivest, R., "The MD5 Message-Digest Algorithm", RFC 1321, April 1992.