Enterprise 네트워크에서 효과적인 정보 알림 시스템 설계

김태범, 박진완, 김명섭
고려대학교 컴퓨터정보학과
{ktb88, jinwan_park, tmskim}@korea.ac.kr

A Design and Implementation of Efficient Information Notifying System on Enterprise Network

Tae-Bum Kim, Jin-Wan Park, Myung-Sup Kim
Dept. of Computer and Information Science, Korea University

요 약

오늘날 대규모 집단이나 단체의 경우 사용자에게 정보를 전달하기 위해 인터넷을 이용한다. 그러나 이러한 인터넷 기반의 알림 시스템의 경우 새로운 정보를 확인하기 위해 사용자가 항상 인터넷을 확인해야 하고, 많은 정보를 제공하는 집단에서는 사용자가 원하는 정보 습득의 어려움을 야기한다. 따라서 본 논문에서는 이러한 문제들을 해결하기 위해 실시간으로 사용자가 원하는 정보를 효과적으로 제공해주는 알림 시스템을 제안하고, 알림 서버가 수집된 정보를 원하는 사용자에게 제공하기까지의 전반적인 처리 과정을 기술한다.

1. 서론

오늘날에는 인터넷의 발달로 많은 사람들이 쉽게 웹에 접근할 수 있는 특성을 이용하여 대규모 집단, 단체에서는 인터넷을 통한 정보제공이 활발하게 이루어지고 있다. 그러나 많은 양의 정보로 인해 정보를 받는 사용자는 원하는 정보를 쉽게 판단할 수 없어 중요한 정보를 제때 받지 못할 수 있다. 그렇기 때문에, 사용자가 원하는 정보를 필터링하여 실시간적으로 알려줄 알림 시스템의 필요성이 부각된다.

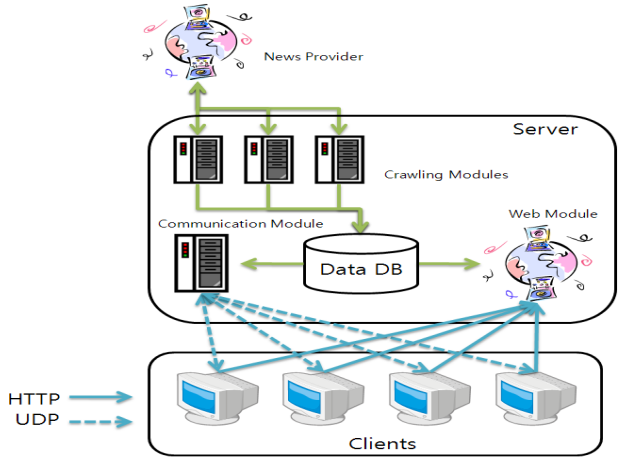
웹 기반의 정보 제공방식 시스템은 정보 제공자가 사용자에게 단방향적인 방식으로 정보 전달이 되고 있다. 이런 단방향 정보제공 시스템의 문제점은 사용자가 새로운 정보를 확인하기 위해 항상 웹에 접속해 확인을 해야 하고, 자신이 필요로 하는 정보를 찾기 위해서는 관련 여러 웹을 검색해야 하는 불편함이 있다. 알림 시스템은 이러한 불편을 해소하고자, 여러 정보 제공 서버의 데이터를 통합 관리하는 DB를 구축하고 사용자마다 원하는 정보를 실시간으로 제공하기 위해 정보 제공자와 사용자의 관계를 서버-클라이언트의 관계로 구성하였다.

본 논문의 2장에서는 앞서 언급한 문제를 해결하기 위해 알림 시스템이 어떤 구조로 이루어져 있는지에 대해 기술한다. 3장에서는 시스템을 이루는 여러 모듈들에 대해 설명하고, 4장에서는 구현 및 성능 평가에 대해 설명하고, 마지막으로 결론 및 향후 과제를 언급하며 본 논문을 마친다.

* 이 논문은 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임.(KRF-2007-331-D00387)

2. 알림 시스템의 구조

(그림 1)은 알림시스템의 전반적인 시스템의 구조이다. 기본적으로 정보 제공자와 사용자간의 관계는 서버-클라이언트의 구조를 갖고 있으며, 서버측은 DB를 중심으로 정보를 가져오는 모듈(Crawling Module), 다수의 클라이언트들과 통신하는 모듈(Communication Module), 사용자가 원하는 정보를 웹으로 보여주는 모듈(Web Module)이 있고, 서버와 통신하는 다수의 클라이언트에는 클라이언트 모듈(Client Module)이 있다.



(그림 1) 알림 시스템의 구조도

2.1. 정보 획득 모듈 (Crawling Module)

클라이언트에게 알릴 정보를 모으는 정보 획득 모듈은 사이트에 의존적인 parsing rule을 갖는 모듈들을 통합 관리하고, 획득한 정보들을 최대한 간단한 정보만을 추려 DB에 저장하는 기능을 담당한다. 이때, DB는 <표 1>과 같이 구성 될 수 있다.

<표 1> DB schema

Type1	Type2	Title	Link
korea.ac.kr	학사 공지	...	http://...
korea.ac.kr	일반 공지	...	http://...

DB는 Type1, Type2, Title, Link등 4개의 entry를 갖으며 Type1로 사이트를 구분하고 Type2로 게시물성격을 구분하며 Title이 해당 게시물의 제목을 나타내며 Link를 통해 정보를 얻은 출처를 나타내며 이 Link를 통해 사용자가 정보의 출처를 확인 할 수 있다. 이 모듈은 결국 클라이언트에게 전달될 정보가 되기 때문에 다양하고 많은 데이터를 저장 할 필요가 있다. 또한, 실시간성이 요구되기 때문에 짧은 주기로 자료를 갱신 할 필요가 있는데 이때, 각 Parsing Module들은 최신의 자료를 판단하여 서버의 부하를 줄일 수 있다.

2.2. 웹 모듈 (Web Module)

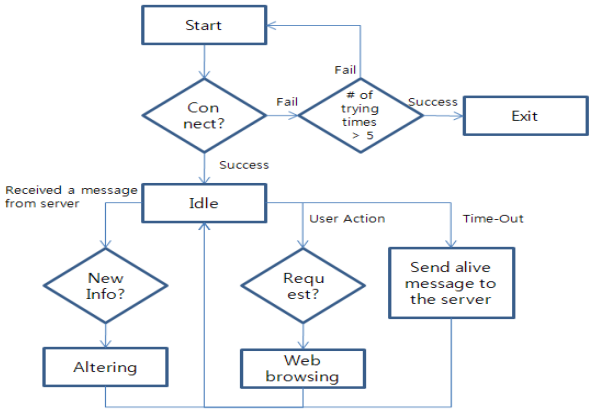
정보를 가져오는 모듈에 의해 DB가 구축이 되면, 해당 정보를 사용자가 원할 때 적절히 보여야 하는데, 이를 담당하는 것이 웹 모듈이다. 서버로부터 전달되는 정보들을 클라이언트 모듈에서 표현하면 불필요한 부하가 생기기 때문에 그러한 부하를 줄이기 위해 사용자가 원하는 정보를 웹으로 표현한다.

2.3. 통신 모듈 (Communication Module)

클라이언트와의 통신을 담당하는 통신 모듈은 클라이언트를 구분할 수 있는 클라이언트의 IP정보, 해당 클라이언트가 어떤 정보를 원하는지에 대한 정보, 그리고 UDP 기반의 통신이므로 지속적으로 Keep-Alive 메시지를 서버로 보내 해당 Client가 살아 있는가를 체크하기 위한 Alive Time 정보를 갖는다. 이 3 가지의 entry로 현재 접속한 클라이언트들의 원하는 정보를 파악하고 정보 획득 모듈에서 새로운 정보를 습득을 알리면 해당 정보를 원하는 클라이언트들에게 알리는 기능을 한다.

2.4. 클라이언트

클라이언트에의 프로그램은 항상 동작하고 있어야 함으로 부하를 최소화 하는 것에 초점을 맞춰야 한다. 클라이언트는 몇 가지 기능으로 이루어져 있다. 첫 번째로는 서버로부터 새로운 정보가 왔는지 체크하는 기능, 두 번째로는 사용자가 특정 정보를 원할 때 웹 모듈로 요청하는 기능, 마지막으로 서버에게 자신이 존재 하고 있다는 것을 알리기 위해 특정 주기로 Keep-Alive 메시지를 보내는 기능이 있다. (그림 2)는 클라이언트 프로그램의 흐름도이다.



(그림 2) 클라이언트 프로그램의 흐름도

3. 구현 및 성능 평가

다음 <표 2>은 서버의 정보 수집 모듈과 통신 모듈의 시간을 측정한 표이다.

<표 2> 모듈 별 실행 시간

종류	시간 (초)
100대의 클라이언트에게 UDP 메시지 전송	$0.96 * 10^{-3}$
20개의 게시판 Crawling 및 Parsing	45

Enterprise Network내의 많은 클라이언트들을 빠르게 관리하는데 있어을 UDP 서버-클라이언트의 관계는 이러한 요구를 만족시키는 것을 확인 할 수 있다. 또한, 20개의 게시판들의 정보 클라에 있어을 대략 45초의 처리 시간이 소요되지만, 단일 프로세싱 처리의 결과로 최악의 시간을 나타낸 것이다. 병렬 처리 Crawling 방식과, 세밀한 Parsing Rule을 적용시킨다면 많은 시간이 단축될 것이다.

4. 결론 및 향후 과제

본 논문에서는 대규모 집단, 단체에서 실시간적으로 중요한 정보를 사용자에게 제공하는데 있어서 알림 시스템의 구조를 제시하였고, 알림 시스템에서의 기능을 통신 모듈(Communication Module), 정보 획득 모듈(Crawling Module), 웹 모듈(Web Module), 클라이언트 모듈(Client Module)로 분류하여 해당 모듈들의 역할들을 살펴보았다.

향후 과제로는 설계한 시스템을 Enterprise 네트워크에 적용시켜 시스템의 효율성을 확인하고자 한다.

참고문헌

[1] 홍석주, 박영배, “역 RSS 웹 크롤링 검색엔진의 설계 및 구현”, 한국통신학회, Apr. 2009, 제 34권 제4호, pp. 139-147.

[2] 박상관, 박건우, 이상훈, “커뮤니티 기반 효율적인 웹 검색 시스템 설계”, 한국정보과학회 학술발표 논문집, Jun. 2008, 제35권 제1호(C), pp. 171-174.

[3] 권 훈, 곽호영, “RSS와 Ajax를 이용한 맞춤형 실시간 정보 서비스 모델 설계”, 한국콘텐츠학회 2007 춘계 종합 학술대회, Jun. 2007, 제5권 제1호, pp. 25-28.