

접근우선순위 부여를 통한 P2P시스템의 신뢰성 향상에 관한 연구

A Study of improving the trustability in P2P System using access priority grant

*# 송현수¹, 문경섭¹

*# H. S. Song(songhyunsu@kifft.re.kr)¹, K. S. Moon¹

¹한국신발피혁연구소

Key words : P2P, access priority, trustability

1. 서론

P2P(Peer to Peer)는 네트워크를 통한 컴퓨터 자원의 공유와 시스템 간의 직접교환이 가능한 서비스로 미국의 파일 공유를 위한 시스템인 Napster를 통해 알려지게 되었다. P2P가 등장하게 된 배경에는 하드웨어의 급속한 발전 및 인터넷의 확산에서 기인한다. 인텔 창시자인 무어(Gorder E. Moore)는 1965년 Electronics Magazine에 반도체 집적도의 발전 속도에 관해 “동일 비용으로 성능이 대략 2년마다 2배 증가한다.” 라는 주된 내용의 이론을 발표한다. 이 이론은 컴퓨터 하드웨어의 급속한 발전을 입증이라도 하듯 하드디스크, 메모리, LCD 화소, 전원장치 등 다양한 분야에 적용되는 사례를 찾을 수 있다. 또한, 인터넷이 널리 보급됨에 따라 활성화된 클라이언트-서버 모델에서는 지나치게 높은 서버의존도로 말미암은 네트워크 과부하, 단방향 커뮤니케이션 등에 따른 문제점이 발생함으로 인하여, 사용자들의 불만을 가져오게 되는 원인이 되었다. 이처럼, 하드웨어의 성능이 급속히 발전함에 따라 서버의 역할을 클라이언트가 대신하게 되고, 인터넷의 보급이 확산됨으로 인해 집중화된 서버에 의존하지 않는 분산 자원의 공유에 대한 관심이 높아지게 되고 이로 인해 P2P가 등장하는 계기가 된다.

P2P 시스템은 운용되는 아키텍처의 형태에 따라 중앙 집중형, 분산형, 혼합형 등으로 분류할 수 있다. 중앙 집중형 P2P 시스템은 피어 명과 콘텐츠 등을 서버에서 관리하는 모델로 Napster, 소리바다 등이 있다. 분산형 P2P 시스템은 각 피어가 서버와 클라이언트의 역할을 하는 것으로 DHT(Distributed Hash Table)를 이용한 Chord, Can Pastry, Tapestry 등이 있으며, 혼합형에는 Super Peer 및 Simple Peer가 존재하는 Kazaa 등이 있다. 이와같이, 아키텍처의 형태 및 응용에 따라 다양한 연구가 진행되고 있지만, P2P 시스템을 사용하는 피어의 신뢰성에 대한 연구는 부족한 실정이다.

일반적인 P2P 파일공유 시스템은 찾고자 하는 파일을 파일명을 통하여 검색한다. 이때 검색된 파일은 검색하고자 했던 파일과 같은 파일명을 가지고 있으나, 파일 내용이 파일명과 일치하지 않는 파일이 검색되거나, 검색된 파일이 실행되지 않거나, 악의적인 목적으로 바이러스가 포함된 파일 등을 유포하는 경우가 발생한다. 위와 같은 문제는 P2P 시스템을 사용할 때 신뢰성을 저하하는 원인이 되며 나아가, 네트워크의 과부하, 컴퓨터의 해킹 혹은 다운 등 치명적인 이차적 문제를 일으키게 된다.

본 논문에서는 사용자의 요청에 의해 검색된 파일에 접근 우선순위를 부여하는 방안에 대해 연구하고, 접근 우선순위를 부여하여 이를 사용자에게 제시함으로써 사용자는 우선순위 목록을 바탕으로 원하는 파일을 다운받을 수 있도록 한다. 이렇게 함으로써, 파일과 파일내용의 불일치, 악의적인 목적의 바이러스 파일 등과 같은 내용의 문제점을 해결하고자 하였다.

2. 본론

2.1 중요도 결정을 위한 중요 요소 결정

접근 우선순위 결정을 위한 사전단계에서는 P2P를 사용함에 있어 중요하다고 판단되는 접속유지시간, 자료의 유효성, 오류, 피어간 거리, 응답속도 등 총 5개의 요소를 결정하였으며, Table 1에 해당 내용을 나타내었다.

Table 1 Definition of each element for access priority

중요 요소	정의	표현
접속 유지시간	24시간을 기준으로 하여 인덱스 서버에 매일 접속한 시간의 평균값	(%)
자료의 유효성	파일명과 파일내용이 다르거나 실행이 되지 않는 파일에 대한 신고 건수	
오류	파일의 찢림, 노이즈, 저질 파일 등 자료의 유효성에 포함되지 않는 모든 오류의 신고 건수	
피어간 거리	전송을 요청하는 피어와 요청받은 자료를 가진 피어 간의 논리적 거리(Hop)	
응답 속도	전송을 요청하는 피어와 요청된 자료를 전송하는 피어 간의 전송 속도	

2.2 AHP를 이용한 중요요소 가중치 결정

본 논문에서는 사용자가 다운로드하고자 하는 파일을 보유한 피어 별 접근 우선순위를 결정하기 AHP기법을 이용하여 중요요소별 중요도를 결정하였으며, 결정된 중요도 및 피어 별 중요요소에 대한 고유 값을 통한 접근 우선순위를 사용자에게 제공하였다. 특정 파일을 보유한 각 피어의 접근 우선순위 결정을 위해 4단계로 구분하여 연구를 진행하였으며, 단계별 내용은 다음과 같다.

① 문제 해결을 위한 계층구조 형성

문제 해결을 위한 계층구조 형성에서는 계층을 총 3개의 계층으로 구성하였으며, 계층 1에서는 접근 우선순위가 결정된 피어의 목록을 제공하고자 하였다. 계층 2에서는 접근 우선순위 결정을 위한 중요요소를 나열하고, 마지막 계층 3에서는 해당 파일을 보유한 다양한 피어를 나열하는 것으로 계층을 형성하였다.

② 설문 및 결과를 토대로 한 쌍대비교 실시

두 번째 단계에서는 중요요소별 중요도 결정을 위해 설문을 설계하였다. 설계된 설문은 결정된 5개의 중요요소에 대해 어떠한 요소가 P2P 파일 공유시스템 이용 시 큰 영향을 미칠지를 요소별 쌍대비교를 통해 결정하도록 하였다. 이때 각 요소별 비교적도는 일반적으로 가장 널리 이용되고 있는 9점 척도를 이용하였으며, 설문에 대한 응답의 거부감을 최소화하기 위해 P2P 파일 공유시스템에 대한 사용 여부, 사용 기간 등 최소한의 범위에서 수집하고 설문을 하였다.

③ 중요요소별 상대적 중요도 계산

2단계에서 설계된 설문을 이용해 P2P 시스템 사용 경험에 있는 사용자를 대상으로 설문을 실시하였으며, Fig 1에 AHP를 적용하여 도출된 중요요소별 상대적 중요도를 나타내었다.

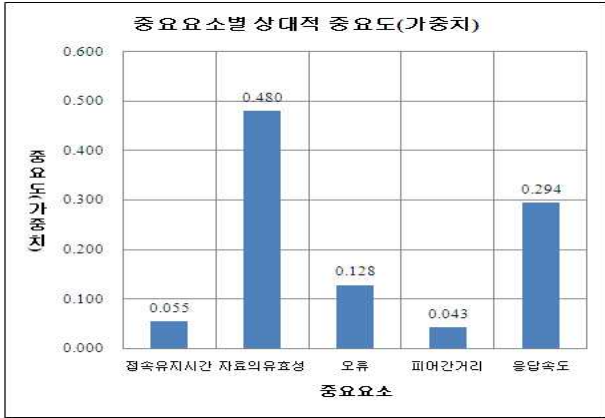


Fig. 1 Relative weight of each element

④ 중요도 통합을 통한 접근 우선순위 목록 제시
 임의의 피어가 특정 파일을 다운로드 받기 위해 검색을 할 때, 3단계에서 결정된 중요요소별 상대적 중요도를 통해 파일별 접근 우선순위 목록을 제시해 준다. 즉, 특정 파일을 가진 임의의 피어에 대한 중요요소 값을 수집하고, 수집된 결과를 상대적 중요도에 적용함으로써, 접근 우선순위 목록을 출력할 수 있도록 하였다.

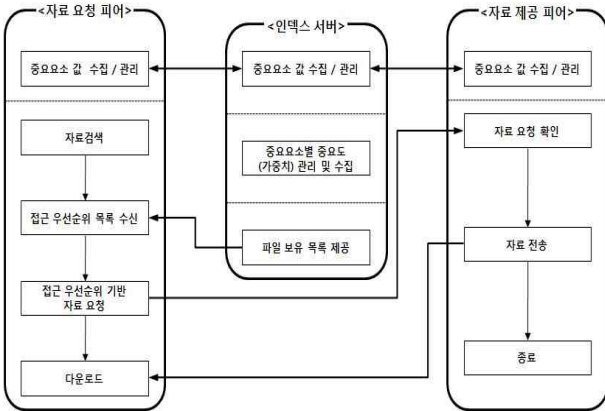


Fig. 2 System operation process

3. 결과

본 논문에서 제안한 접근 우선순위 기반 P2P 시스템의 평가는 파일 집단 중 파일명과 파일내용이 동일한 파일을 제공해 주는 정도를 통해 성능 평가를 하고자 하였으며, 비교 대상은 현재 서비스 중인 P2P Client(dEonkey, Bittorrent, 소리바다)에서 제공해 주는 파일명, 파일크기, 재생시간, 음질, 전송속도 등의 항목 중 전송속도를 기준으로 제공해 주는 목록을 비교 대상으로 결정하였다.

전체 파일 중 참인 파일 비율은 유료서비스 P2P 시스템과 무료 P2P 시스템에 따라 차이가 있으나, 전체 파일 중 30%, 50%, 70%가 참인 파일이 될 수 있도록 각각의 파일 그룹을 생성하였다.

전체 파일 중 참인 파일의 비율이 약 30%인 파일 집단에서, 파일명에 따라 응답속도 기준 검색과 접근 우선순위 검색 결과를 Fig 3에 나타내었다. 검색결과에서 상위 10개 파일 중 참인 파일의 개수는, 응답 속도 기준 검색 결과와 비교하면 본 논문에서 제시한 검색 방법에 대한 결과가 0~4개 정도 많은 참인 파일을 제시해 주는 것으로 확인되었다.

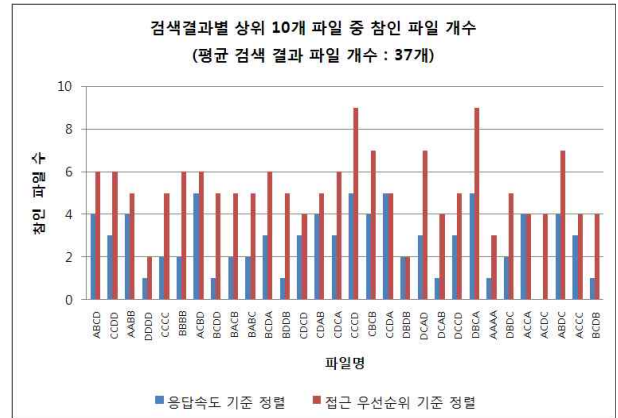


Fig. 3 Search result of file group in 30% truth

Fig 4는 파일 집단의 성능 향상 정도를 나타낸 것으로, 전체 파일에서 참인 비율이 30%인 파일 집단에서는 응답속도 기준 검색보다 본 논문에서 제안한 검색 방법이 24.3% 정도 더 많은 참인 파일을 제공해 주는 것으로 나타났으며, 50% 그룹에서는 약 19.7%, 70% 그룹에서는 약 12% 정도 더 많은 참인 파일을 제공해 주는 것으로 나타났다. 이와 같이, 참인 파일 비율에 따른 차이는 있으나, 본 논문에서 제시한 검색 방법이 응답속도 기준 검색보다 성능이 향상된 것으로 나타났다. 또한, 전체 파일 중 참인 파일의 비율이 줄어들수록 본 논문에서 제시한 검색 방법이 참인 파일을 제시해 주는 성능이 뛰어난 것으로 확인되었다.

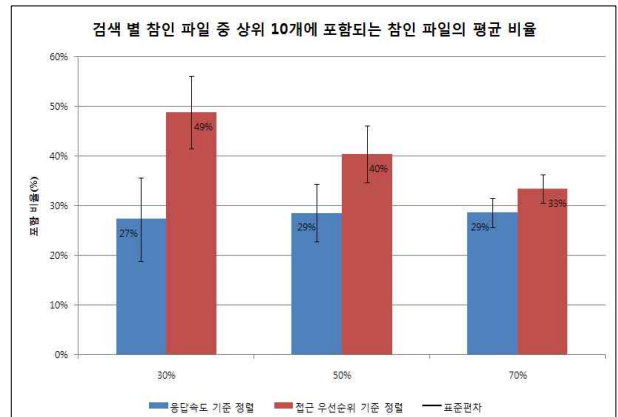


Fig. 4 Average ratio of truth file in top 10 search result

4. 결론

본 논문에서는 P2P 파일 공유시스템에 참여하는 피어에 대한 중요요소 값과 중요요소별 중요도를 결정해 각 피어가 가진 파일별 우선순위를 부여하고, 사용자에게 접근 우선순위 목록을 제시함으로써 사용자로 하여금 신뢰성이 높은 파일을 다운로드 받을 수 있는 방안을 제안하였다. 향후 연구에서는 콘텐츠 종류에 따라 다른 중요도를 고려하는 부분에 대한 연구를 진행함과 동시에 시스템의 배포를 통해 제안 방식의 정확한 성능 평가가 이루어져야 할 것으로 사료된다.

참고문헌

- DS Milojicic, V Kalogeraki, R Lukose, K Nagaraja, "peer to peer computing", Tech Report: HPL-2002-57, HP, July 2003
- J Liang, R Kumar, Y Xi, KW Rose, "Pollution in p2p file sharing systems", Proceedings IEEE INFOCOM 2005. 24th Annual Joint Conference, 2005