

모바일 P2P 네트워크에서 효율적인 콘텐츠 검색을 위한 데이터 배포 기법

Data Dissemination Method for Efficient Contents Search in Mobile P2P Networks

복경수, 조미림, 유재수

충북대학교 정보통신공학과 및 충북BIT연구중심대학육성사업단

Kyoung-Soo Bok(ksbok@chungbuk.ac.kr), Mi-Rim Cho(mirim3441@gmail.com),
Jae-Soo Yoo(yjs@chungbuk.ac.kr)

요약

모바일 P2P 네트워크를 위해 제안된 데이터 배포 기법들은 프로파일과 일치하는 콘텐츠 검색 성능은 매우 뛰어나지만 프로파일과 일치하지 않는 콘텐츠 검색의 경우 질의 처리를 위한 추가적인 비용이 발생하기 때문에 프로파일과 일치하지 않는 콘텐츠 검색 성능 향상에 대한 추가적인 고려가 필요하다. 이러한 문제를 해결하기 위해 본 논문에서는 모바일 P2P 환경에서 효율적인 콘텐츠 검색을 위한 새로운 데이터 배포 기법을 제안한다. 제안하는 기법에서는 타임스탬프 메시지를 사용함으로써 이전 통신 경험 여부를 판단하고 이에 따른 데이터 배포를 수행한다. 또한, 제한된 메모리에 배포된 데이터를 효율적으로 저장하기 위해 랭킹 기법을 제안한다. 제안하는 랭킹 기법은 프로파일 일치여부 뿐만 아니라 주변의 배포 범위, 데이터를 배포해준 피어와의 연결성을 고려함으로써 차후의 질의 배포를 감소시킬 수 있다.

■ 중심어 : | 모바일 P2P | 데이터 배포 | 콘텐츠 검색 | 랭킹 |

Abstract

In the existing data dissemination methods for mobile P2P networks, the search performance of content that matches the peer profile is very excellent. However, in the search for content that does not match the their profile, additional consideration about case that contents does not match the profile is needed because costs for the query processing will be incurred. To solve these problems, we propose a new data dissemination method for efficient contents search in mobile P2P networks. In the proposed method, peers determine whether they experienced communications by using a timestamp message and then perform data dissemination. We also propose a ranking algorithm to efficiently store dissemination data in a limited memory. The proposed ranking method can reduce query messages by considering the profile matches, the distribution range, and the connectivity to the data distribution peer.

■ keyword : | Mobile P2P | Data Dissemination | Content Search | Ranking |

* 이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 산학협력 선도대학(LINC)육성사업(No. 2012-B-0013-010112)과 KISTI의 지원(K-12-L06-C02-S03)으로 수행되었음

접수번호 : #120709-005

심사완료일 : 2012년 08월 07일

접수일자 : 2012년 07월 09일

교신저자 : 유재수, e-mail : yjs@chungbuk.ac.kr

I. 서론

최근 스마트폰이나 PDA, 태블릿 PC와 같은 이동 기기의 발달에 따라 기존의 P2P 시스템에 이동성이 결합된 MP2P 네트워크(Mobile Peer-to-Peer Network)에 대한 연구가 진행되고 있다. MP2P 네트워크는 IEEE 802.11, 블루투스(Bluetooth)와 같은 단거리 무선 통신을 사용하는 이동 기기들로 구성되어 있다[1][2]. 이러한 모바일 P2P는 소셜 네트워크[3], 파일 공유[4][5], 광고 및 추천[6][7] 등의 응용에 사용되고 있다. 신뢰할만한 높은 연결성을 보장하는 유선 기반의 P2P 시스템과 달리 MP2P 시스템에서는 각각의 피어들이 제한된 대역폭, 배터리, 컴퓨팅 능력, 저장 공간을 가지고 있다. 또한, 피어의 이동성으로 인해 네트워크 토폴로지가 자주 변경되며 이로 인해 유선 환경의 P2P 시스템과 같이 완벽하게 연결성을 보장하기 어렵다[8-10].

넓은 통신 대역폭과 높은 전송 속도, 충분한 저장 공간을 갖는 유선 환경의 P2P 응용에서는 전체 네트워크에 데이터 즉, 콘텐츠에 대한 메타데이터를 배포하고 이를 통해 콘텐츠를 검색 및 요청할 수 있도록 하였다. 그러나 모바일 P2P 환경에서는 이동 기기들의 제약 사항 즉, 제한된 메모리, 배터리 성능, 잦은 토폴로지 변경으로 인해 기존의 P2P 응용과 같이 전체 네트워크에 데이터를 배포하는 것이 어렵다[11]. 따라서 모바일 P2P 응용에 적합한 새로운 데이터 배포 기법이 필요하다. 현재까지 연구된 모바일 P2P 응용은 피어들의 정보를 이웃한 피어들에게 전송하고 질의를 처리하는 방식을 사용하고 있다[12].

B. Xu, F. Vafaee, O. Wolfson[13] 등은 다수 콘텐츠에 대한 메타데이터를 질의와 함께 배포함으로써 피어들에 존재하는 특정 콘텐츠의 개수(공급)와 콘텐츠가 질의로부터 호출되는 즉, 질의의 결과가 되는 횟수(수요)를 이용한 마켓 알고리즘을 사용한다. 마켓 알고리즘을 이용하여 각각의 데이터들의 우선 순위를 결정하고 이것을 바탕으로 인덱스를 저장 및 배포한다. B. Qureshi, G. Min, D. D. Kouvatso, M. Ilyas[14] 등은 모바일 P2P 네트워크에 소셜 개념을 적용하기 위해 주기적으로 프로파일 즉, 관심사를 포함하는 메시지를 브

로드캐스팅한다. 이 메시지를 받은 피어들은 자신과 관심사가 동일할 경우 메시지를 전송한 피어의 관심사와 일치하는 속성을 갖는 콘텐츠를 전달한다. 그러나 [13]에서 제안된 기법의 경우 데이터 배포에 따른 질의 배포 비용이 추가로 발생한다. [14]는 프로파일을 기반으로 하고 있기 때문에 자신의 관심사와 일치하는 콘텐츠의 검색 성능 향상에만 초점을 맞추고 있다.

본 논문에서는 제한된 메모리를 효율적으로 사용하고 사용자의 관심사와 일치하는 콘텐츠에 대한 검색 성능뿐만 아니라 자신의 프로파일과 일치하지 않는 콘텐츠 검색에 있어서도 높은 성능을 발휘할 수 있도록 새로운 데이터 배포 기법을 제안한다. 제안하는 데이터 배포 기법은 제약적인 네트워크와 메모리를 사용하는 피어에 콘텐츠 검색에 필요한 정보를 최대한 저장하고 전송하기 위해 사용자 관심사 뿐만 아니라 데이터 배포 정도 및 피어와의 연결성 등을 고려한다. 또한, 타임스탬프 메시지를 통해 피어간의 통신 경험 여부를 판단하고 이를 이용해 차후에 발생할 수 있는 중복적인 데이터의 전송을 제거하였다. 이러한 데이터 배포는 네트워크에 질의를 전달하는 횟수를 감소시키고 검색 성능을 향상시킬 수 있다.

본 논문의 나머지 구성은 다음과 같다. II장에서는 기존에 제안된 데이터 배포 기법과 기존 연구의 문제점을 기술한다. III장에서는 모바일 P2P 환경에서 효율적인 콘텐츠 검색을 위한 데이터 배포 기법을 기술하고 IV장에서는 기존에 제안된 데이터 배포 기법과의 비교 분석을 통해 제안하는 기법을 우수성을 입증한다. 마지막으로 V장에서는 결론과 향후 연구 방향을 제시한다.

II. 관련 연구

B. Xu 등[13]은 서버가 없는 환경에서 이동 기기들만으로 질의를 처리하기 위한 마켓 알고리즘을 제안하였다. 마켓 알고리즘은 데이터 배포 과정에서 질의를 함께 배포함으로써 모바일 피어들에 존재하는 특정 콘텐츠의 개수(공급)와 이 콘텐츠가 질의로부터 호출되는 횟수(수요)를 이용하여 콘텐츠들의 우선 순위를 결정한다.

다. 피어는 새로운 피어를 만나면 자신이 소유하고 있는 콘텐츠들의 식별자와 질의 데이터베이스에 저장된 질의를 새로운 피어에게 전송한다. 이를 통해 새로운 피어로부터 자신만이 소유하고 콘텐츠 식별자와 새로운 피어만이 소유하고 있는 콘텐츠 식별자를 전송받는다. 그 후 피어는 새롭게 갱신된 콘텐츠 정보만을 이웃 피어들에게 전달한다. 이 과정에서 다른 피어로부터 받은 정보를 제한된 메모리에 저장하기 위해 마켓 알고리즘을 사용한다. 그러나 이 기법은 기본적으로 질의 배포를 위한 비용이 발생하고, 수요와 공급에 따라 저장할 데이터의 우선 순위를 결정하기 때문에 수요가 낮은 즉, 검색 빈도가 적은 콘텐츠 검색에 있어서 성능이 저하된다.

B. Qureshi 등[14]는 프로파일 즉, 관심사와 일치하는 콘텐츠에 대한 정보를 배포하기 위한 기법을 제안하였다. 이 기법에서 피어들은 주기적으로 자신의 프로파일을 포함하는 Announce 메시지를 브로드캐스팅한다. 이 메시지를 받은 주변의 피어들은 자신의 콘텐츠 리스트에서 메시지에 포함된 관심사와 일치하는 속성을 갖는 콘텐츠들에 대한 인덱스 만들어 Invite 메시지에 넣어 전송한다. Announce 메시지를 보낸 피어는 수신한 Invite 메시지를 확인하고 원하는 콘텐츠가 있는 경우 Request 메시지에 콘텐츠의 식별자를 보냄으로써 해당 파일을 요청한다. 마지막으로 콘텐츠를 소유하고 있는 피어는 Send 메시지에 실제 콘텐츠를 담아 전송한다. 이 기법은 소셜 네트워크 환경에서 동일한 관심사를 가지는 사용자들을 관계성을 가지고 연결해주고, 자신의 프로파일과 일치하는 콘텐츠에 대한 검색 시 매우 좋은 성능을 나타낸다. 그러나 자신의 프로파일과 일치하지 않지만 대중적으로 인기 있는 콘텐츠에 대한 검색 시에는 추가적인 질의 배포 비용이 발생한다.

III. 제안하는 데이터 배포 기법

1. 시스템 개요

본 논문에서는 최대한 이웃 피어에 질의 전달하지 않고 이동하는 피어들이 원하는 콘텐츠를 검색하기 위한

데이터 배포 기법을 제안한다. 이전에 연구된 데이터 배포 기법들은 피어간의 통신 경험을 고려하지 않아 중복적인 데이터 배포가 발생한다. 또한, 프로파일 및 대중성을 고려하여 데이터를 저장하기 때문에 프로파일과 일치하지 않는 콘텐츠를 검색하거나 대중적이지 않은 콘텐츠를 검색할 때 성능이 저하되는 문제점이 있다. 따라서, 제안하는 기법에서는 데이터 중복 배포 문제, 데이터를 저장하기 위한 랭킹 기법을 제안한다.

각 피어들은 자신의 로컬에 콘텐츠가 추가, 삭제되거나 다른 피어로부터 받은 데이터들로 인해 변경 사항이 발생할 때에만 통신을 수행한다. 각 피어는 주변 피어들에게 타임스탬프 메시지를 배포하고 타임스탬프 메시지를 이용해 통신이력을 확인한다. 실제 데이터 배포는 통신 및 데이터 변경 유무에 따라 결정된다. 제안하는 기법에서는 두 피어간의 정보 교환 시, 이전에 통신한 경험 여부를 확인하고 통신 경험이 없는 경우 전체 정보를 모두 전송한다. 그러나 통신 경험이 있는 경우에는 이전 통신을 수행한 이후 변경된 정보만을 전달한다.

피어는 프로파일 정보, 로컬 인덱스, 원격 인덱스, 피어 정보 테이블을 유지한다. 프로파일의 경우 사용자가 선택한 관심항목을 등록함으로써 설정된다. 기본적으로 각 피어들은 다른 피어를 만나면 타임스탬프 메시지를 전송한다. 이 타임스탬프 메시지는 두 피어간의 연결가능 시간과 이전에 두 피어 간에 통신 경험이 있는지 확인하기 위해 사용된다. 타임스탬프 메시지는 {PID, LUT, PVI}으로 구성되어 있다. 이때, PID는 자신의 식별자를, LUT(Last Update Time)는 자신의 정보가 마지막으로 변경된 시간, PVI(Position Vector Information)는 자신의 이동 방향 및 속도를 포함하는 벡터 정보를 나타낸다. 타임스탬프 메시지는 피어 정보 테이블을 이용해 이전에 해당 피어와 통신한 경험이 있는지 판단하는데 사용된다. 피어 정보 테이블은 {PID, LCT, PVI}로 구성되어 있다. 이때, 각각 PID는 피어의 식별자, LCT(Last Communication Time)는 해당 피어와 마지막으로 통신한 시간, PVI는 그 시점에 해당 피어의 벡터 정보를 나타낸다.

다른 피어로부터 타임스탬프 메시지를 받은 피어는

우선 메시지에 포함된 PID가 피어 정보 테이블에 있는지 확인한다. 만약 피어 정보 테이블에 해당 식별자가 있다면 이전에 통신 경험이 있는 피어로 판단하고 데이터 배포 시 피어와의 마지막 통신 시간인 LCT 이후의 변경 정보만을 해당 피어에게 전달한다. 그러나 피어 정보 테이블에 해당 식별자가 없다면 처음 통신하는 피어로 판단하고 자신이 소유한 모든 데이터를 해당 피어에게 전달한다.

각 피어는 데이터 배포를 위해 로컬 인덱스와 원격 인덱스를 유지한다. 로컬 인덱스는 실제 피어가 소유하고 있는 콘텐츠에 대한 인덱스를, 원격 인덱스는 다른 피어들로부터 수집된 콘텐츠의 인덱스들을 저장하고 있다. 로컬 인덱스는 {CID, CT}로 구성되어 있다. 이때, CID(Content ID)는 콘텐츠 식별자, CT(Content Type)는 콘텐츠 유형을 나타낸다. 로컬 인덱스는 자신이 소유한 콘텐츠들에 대한 정보를 유지하기 때문에 별도로 콘텐츠를 소유자에 대한 정보를 저장하지 않는다. 이와 달리 원격 인덱스의 경우 {CID, PID, CT, POO}로 구성되어 있으며 해당 콘텐츠를 소유한 피어의 식별자 PID와 콘텐츠를 실제 소유한 해당 피어 정보를 연결하는 POO(Pointer Of Owner)를 포함한다. 각 피어들은 통신을 통해 수집한 피어 정보를 피어 정보 테이블에 저장하기 때문에 POO는 피어 정보 테이블에 저장된 피어 정보를 가리키는 포인터이다. [그림 1]은 원격 인덱스와 피어 정보 테이블을 나타낸 것이다.

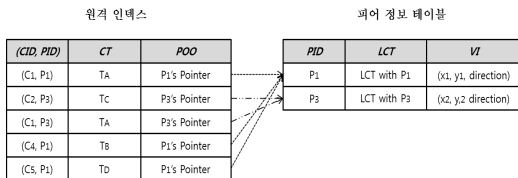


그림 1. 원격 인덱스와 피어 정보 테이블

2. 데이터 배포 과정

각 피어는 자신의 로컬 콘텐츠가 변경되거나 원격 인덱스가 변경되었을 때 데이터 배포를 수행한다. 먼저 타임스탬프 메시지를 주변 피어들에게 전송하고 각 피어들은 수신한 타임스탬프 메시지와 자신의 피어 정보

테이블을 이용하여 주변 피어와의 통신 경험을 판단한다. 통신 경험의 유무에 따라 전달할 정보의 내용을 결정하고 타임스탬프 메시지를 전송한 피어에게 콘텐츠 정보를 전송한다. 주변 피어들로부터 정보를 수신한 피어는 랭킹 기법을 통해 자신의 원격 인덱스를 갱신하고 변경된 내용을 주변의 피어들에게 재배포한다.

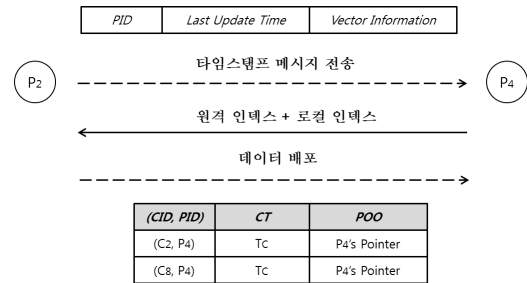


그림 2. 제안하는 데이터 배포 기법의 예

[그림 2]는 제안하는 데이터 배포 기법의 예를 나타낸다. P2가 이동하는 과정에서는 P4와 통신이 되었다고 가정하자. 데이터를 배포하기 전에 P2는 P4에게 타임스탬프 메시지를 전송한다. P2의 타임스탬프 메시지를 수신한 P4는 우선 P2의 벡터 정보와 자신의 벡터 정보를 이용해 통신 가능 시간을 계산해 최소 통신 시간 이상 연결이 가능한지를 판단한다. 이때, 최소 통신 시간은 피어들 사이에 타임스탬프 메시지를 상호 교환하고 실제 데이터를 전송하는데 소요되는 시간을 의미한다. 최소 통신 시간 이상을 만족하지 못하면 P4는 P2의 타임스탬프 메시지를 무시하고 P2와 데이터를 교환하지 않는다. 그러나 최소 통신 시간을 만족하는 경우 P4는 자신의 피어 정보 테이블에 P2의 식별자가 존재하는지 확인함으로써 기존에 통신한 경험이 있는지를 확인한다. 만약 통신이력이 없고 새롭게 통신을 한 피어인 경우 자신의 로컬 인덱스에 저장된 데이터와 원격 인덱스에 저장된 데이터 모두를 P2에게 전송한다. 만약 통신 이력이 있다면 P2와 통신을 수행한 마지막 시간 이후 변경된 정보만을 전송한다. P2는 원격 인덱스에 수신한 정보를 저장하기 위해 P4로부터 받은 데이터들과 이전에 원격 인덱스에 저장되어 있던 데이터들에 대한 우선 순위를 결정한다. 결정된 순위에 따라 P2는 자신

의 원격 인덱스를 갱신하고, P4가 소유한 데이터를 제외된 데이터들을 이웃 피어들에게 브로드캐스팅함으로써 P4로 인해 갱신된 P2의 정보를 다른 피어들에게 전파한다.

(CID, PID)	CT	POO
(C1, P1)	TA	P1's Pointer
(C2, P3)	TC	P3's Pointer
(C1, P3)	TA	P3's Pointer
(C4, P1)	T8	P1's Pointer
(C8, P1)	T8	P1's Pointer

(a) 원격 인덱스

(CID, PID)	CT	POO
(C2, P6)	Tc	P6's Pointer
(C7, P5)	TA	P5's Pointer
(C8, P4)	Tc	P4's Pointer

(b) 수신한 데이터

(CID, PID)	CT	POO
(C1, P1)	TA	P1's Pointer
(C2, P3)	Tc	P3's Pointer
(C2, P6)	Tc	P6's Pointer
(C7, P5)	TA	P5's Pointer
(C4, P1)	Tc	P1's Pointer

(c) 갱신된 원격 인덱스

그림 3. 피어 P2의 우선 순위 결정에 따른 원격 인덱스 갱신 예

[그림 3]은 피어 P2의 원격 인덱스 갱신 과정을 나타낸 것이다. 피어 P2가 이전에 피어 P4와 통신한 경험이 있다면 피어 정보 테이블에 저장된 피어 P2와의 마지막 통신 시간을 확인한다. 그 후 P4의 데이터 중에서 P2와의 마지막 통신 시간 이후에 변경된 데이터 즉, (C2, P6), (C7, P5), (C8, P4)만을 P2에게 전송한다. 변경된 데이터가 없다면 P2에게 아무런 데이터도 전송하지 않는다. P2는 P4로부터 전달받은 {(C2, P6), (C7, P5), (C8, P4)} 데이터와 자신의 원격 인덱스에 저장된 모든 데이터들에 대해 새롭게 순위를 계산한다. [그림 3]의 (c)에서 보는 것과 같이 전송받은 {(C2, P6), (C7, P5), (C8, P4)} 중에서 (C8, P4)를 제외된 {(C2, P6), (C7, P5)} 데이터만이 반영된 것을 확인할 수 있다. 데이터 배포 과정의 마지막에는 주변으로부터 수신한 {(C2, P6), (C7, P5), (C8, P4)}를 주변 노드들에게 브로드캐스팅 함으로써 주변 노드들에게 전달한다.

3. 우선 순위 기반의 데이터 저장

각 피어는 데이터 배포를 통해 네트워크에 존재하는 피어들이 소유한 콘텐츠들의 정보를 알고 있으므로써

콘텐츠 검색 성능을 향상시킬 수 있다. 그러나 주변으로부터 수신한 모든 데이터들을 제한된 메모리 내에 모두 저장하는 것은 불가능하다. 따라서 각 피어는 수신한 데이터들 중에서 어떠한 정보를 저장할 것인지를 결정해야 한다. 피어가 요청하는 정보를 최대한 이웃 피어에서 결과를 얻을 수 있도록 하기 위해 우선 순위 결정을 위한 랭킹 기법을 제안한다.

P2P 응용에서는 모든 사용자가 자신의 프로파일과 정확하게 일치하는 콘텐츠만을 검색하지는 않는다. 예를 들어, 영화 파일을 하나의 콘텐츠로 봤을 때 어떤 사용자의 관심사(장르)가 공포 영화라고 해서 오직 공포 영화만을 검색하지는 않는다는 것이다. 액션 영화가 새로 개봉했고 많은 사람들이 그 영화에 관심을 갖는다면 공포 영화를 좋아하는 사용자도 액션 영화에 대해 관심을 가지고 해당 영화 파일을 얻기를 원할 수도 있다는 것이다. 그렇기 때문에 우선 순위를 통해 데이터를 저장하면 이전의 프로파일(관심사) 일치 여부만으로 데이터 저장 여부를 결정하던 기법에 비해 자신의 관심사와 일치하는 콘텐츠에 대한 검색 성능뿐 아니라 자신의 관심사와는 일치하지 않더라도 많은 사용자들이 관심을 가지는 콘텐츠에 대한 검색 성능도 향상시킬 수 있다.

대부분의 사용자는 많은 사람들이 관심을 갖는 콘텐츠 보다는 자신의 관심사와 일치하는 콘텐츠를 검색하는 경우가 더 많기 때문에 데이터의 우선 순위를 결정할 때 콘텐츠의 속성이 피어가 가지고 있는 프로파일과 일치하는가의 여부를 최우선적으로 고려한다. 수신된 데이터의 콘텐츠 속성이 자신의 프로파일 항목과 일치한다면, 해당 데이터의 관심사 일치 여부 값 (*match_value*)은 1의 값을 갖고 일치하지 않는다면 0의 값을 갖는다.

주변에 얼마나 많은 피어들이 해당 콘텐츠에 대한 데이터를 저장하고 있는지를 나타내는 데이터 배포 범위 또는 데이터 배포 수를 우선순위 결정에 고려한다. 다른 피어로부터 데이터를 수신한 피어의 입장에서는 해당 콘텐츠에 대한 데이터를 주변의 피어들이 많이 가지고 있을 경우 보다는 적게 가지고 있는 데이터를 저장하는 것이 더 효율적이다. 이것은 차후에 콘텐츠 검색을 고려한 것으로 검색하고자 하는 콘텐츠에 대한 정보

를 로컬에 검색하지 못하는 경우에 주변의 피어들에게 질의를 배포해 데이터를 검색해야 한다. 이 경우 주변에 피어들이 해당 데이터를 많이 저장하고 있는 경우는 멀리까지 질의를 배포하지 않아도 원하는 결과를 빠르게 얻을 수 있지만 주변 피어들이 해당 데이터를 저장하고 있지 않다면 원하는 결과를 얻기 위해 질의를 멀리까지 배포해야만 한다. 이때, 자신이 해당 데이터를 저장하고 있다면 주변 피어들의 해당 콘텐츠 검색 시에 중계자 역할을 함으로써 피어들의 검색 성능을 향상시킬 수 있기 때문이다. 콘텐츠 C_i 의 데이터 배포 정도 ($dist_rate$)는 (식 1)을 사용해 계산한다. 이때, $owner_num(C_i)$ 은 콘텐츠 C_i 의 소유자 정보에 등록된 피어의 수이고, max_owner_num 은 하나의 콘텐츠에 소유자 정보로 등록될 수 있는 최대 피어의 수를 나타낸다. 소유자 정보에 등록된 피어의 수가 등록 가능한 최대 피어의 수에 가까울수록 0에 가까운 값을 가지게 된다. 즉, 주변에 콘텐츠 C_i 에 대한 데이터가 적게 배포되어 있을수록 dis_rate 는 1에 가까운 값을 갖는다. $match_value$ 와 dis_rate 을 이용해 데이터의 우선 순위를 결정하는 $priority(C_i)$ 는 (식 2)과 같다.

$$dis_rate(C_i) = 1 - \frac{owner_num(C_i)}{max_owner_num} \quad (1)$$

$$priority(C_i) = \frac{match_value(C_i) + dis_rate(C_i)}{2} \quad (2)$$

콘텐츠에 대한 데이터를 전송해 준 피어와의 연결성 ($connect_value$)을 고려한다. 각 피어는 타임스탬프 메시지를 통해 전달받은 피어의 백터 정보와 자신의 백터 정보를 이용해 두 피어간의 연결성을 계산한다. 제안하는 기법에서는 고정된 n-홉 이내 통신이 가능한 시간을 연결성 판단 기준으로 정의한다. 연결성을 계산하는 방법은 배포 정도를 계산하는 것과 유사하다. (식 3)에서 $connection_time(N_i, N_j)$ 는 피어 N_i 와 N_j 사이에 n-홉 이내의 통신이 가능한 시간을 의미하고 $max_connection_time(N_i)$ 는 계산 시점에 피어 N_i 의 입장에서 다른 피어들과의 가장 긴 $connection_time$ 을 의미한다.

$connect_value(N_i, N_j)$ 는 각 피어 기준으로 n-홉 이내의 통신 가능 시간이 길수록 1에 가까운 값을 갖는다.

$$connect_value(N_i, N_j) = \frac{connection_time(N_i, N_j)}{max_connection_time} \quad (3)$$

최종적으로 저장 여부를 결정하기 위한 콘텐츠 C_i 에 대한 우선 순위는 (식 4)을 통해 계산된다. 이때, α 와 β 각각은 사용자에게 의해 결정되며 $\alpha + \beta = 1$ 이다. 만약 α 값이 더 큰 경우 콘텐츠를 소유한 피어와의 연결성보다는 프로파일 일치여부와 대중성에 더 큰 비중을 두게 되고 β 값이 더 큰 경우는 그와 반대의 비중을 두게 된다. 콘텐츠 C_i 에 대한 순위는 C_i 의 소유자 정보에 등록된 피어들에 대해 각각 된다. 이것은 하나의 콘텐츠 데이터에 대해 등록할 수 있는 최대 피어의 수가 정해져 있기 때문에 연결성이 좋은 피어들을 우선적으로 소유자로서 등록하기 위함이다.

$$rank(C_i) = \alpha(priority(C_i)) + \beta(connect_value(N_i, N_j)) \quad (4)$$

IV. 성능평가

제안하는 기법의 우수성을 보이기 위해서 [14]에서 제안한 기법과 시뮬레이션을 통해 성능 평가를 수행한다. 성능 평가는 [표 1]과 같은 환경에서 수행하였다. 성능 평가를 수행하기 위해 Intel Core i3 3.07GHz CPU와 4G RAM을 사용하는 컴퓨터에서 JDK 1.7을 사용하여 자바로 구현하였다. 제안하는 기법의 우수성을 보이기 위해 피어의 수와 피어의 이동성에 따라 데이터 배포 비용, 배포되는 데이터 크기, 검색 성능을 평가하였다. 이동 속도에 따른 성능 평가를 수행하기 위해 [15]에서 제공하는 Network-based Generator of Moving Objects를 이용하여 피어의 이동 좌표와 속도를 추출하였다. 이동 속도는 fast, middle, slow로 구분하였으며 max.speed div 값을 각각 10, 50, 250으로 설정하였다.

표 1. 성능 평가 파라미터

구분	내용
실험 크기	100m X 100m
통신 범위	20m
피어 수	50 ~ 100

모바일 P2P 환경에서는 실제 데이터에 대한 정확성도 중요하지만 피어들이 제한된 대역폭을 사용하기 때문에 데이터 배포를 위한 메시지 송수신이 네트워크 전반에 영향을 미칠 수 있다. 따라서 데이터 배포를 위해 발생하는 메시지 수를 측정함으로써 네트워크에 미치는 영향력을 확인하였다. [그림 4]는 피어의 이동 속도에 따른 데이터 배포 비용에 대한 성능 평가 결과를 나타낸 것이다. 데이터 배포 비용은 데이터 배포를 위해서 발생하는 메시지수를 측정함으로써 평가하였다. 제안하는 데이터 배포 기법은 전반적으로 기존 기법에 비해 우수한 성능을 나타낸다. 제안하는 기법은 기존 기법에 비해 피어의 수가 50일 때에는 약 10배, 피어의 수가 100일 때에는 약 25배 많은 메시지가 발생하는 것을 확인할 수 있다. 제안하는 기법은 전달받은 배포 메시지를 자신의 이웃 피어들에게 계속해서 재전송하는 비용이 추가적으로 발생한다.

제안하는 기법에서는 불필요한 중복적인 데이터의 통신을 제거하기 위해 타임스탬프 메시지를 통해 통신 경험을 비교한 후 데이터 배포를 수행한다. 타임스탬프 메시지 사용을 통해 실제로 중복적인 데이터 전송이 제거되고 그에 따른 네트워크 전반에 걸쳐 배포되는 데이터의 감소를 확인하기 위해 배포되는 데이터 크기를 측정하였다. [그림 5]는 배포되는 데이터의 크기를 나타낸다. 배포되는 데이터의 크기는 100 라운드 동안 데이터 배포를 수행하면서 전체 네트워크에 배포되는 패킷의 수로 측정하였다. 하나의 패킷은 하나의 콘텐츠에 대한 메타 데이터를 나타낸다. [그림 5]에 나타난 것처럼 제안하는 기법은 기존 기법에 비해 배포되는 데이터의 크기는 크게 감소하는 것임을 알 수 있다. 기존 기법은 피어의 수가 50일 때 2배 이상, 피어의 수가 100일 때 4배 이상 배포되는 데이터 크기를 큰 것을 확인할 수 있다. 제안하는 기법의 경우 피어들과의 통신 이력을 저장함으로써 동일한 피어에게 중복적인 데이터를 전달하지 않는다. 기존 기법은 자신이 소유한 모든 콘텐츠들의 데이터를 배포하지만 제안하는 기법에서는 이전의 통신 경험을 비교해 변경된 정보만을 배포함으로써 상대적으로 더 적은 크기의 데이터를 배포하게 된다.

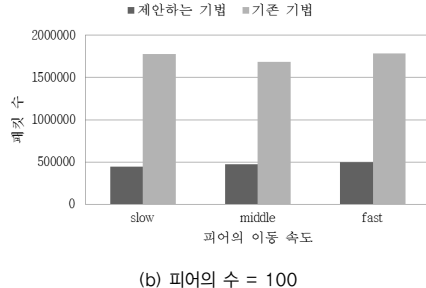
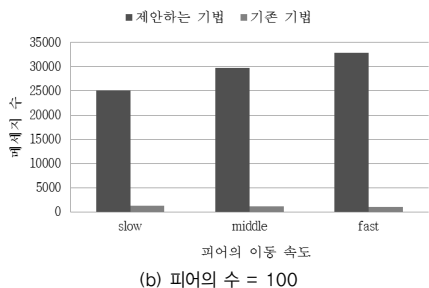
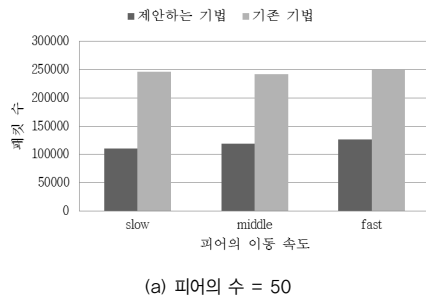
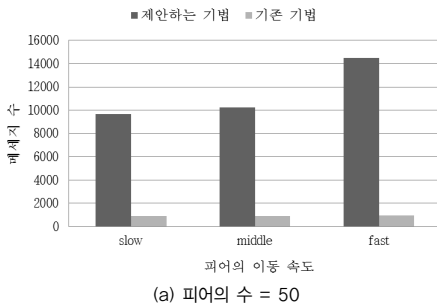


그림 4. 데이터 배포를 위한 메시지 수

그림 5. 배포되는 데이터 크기

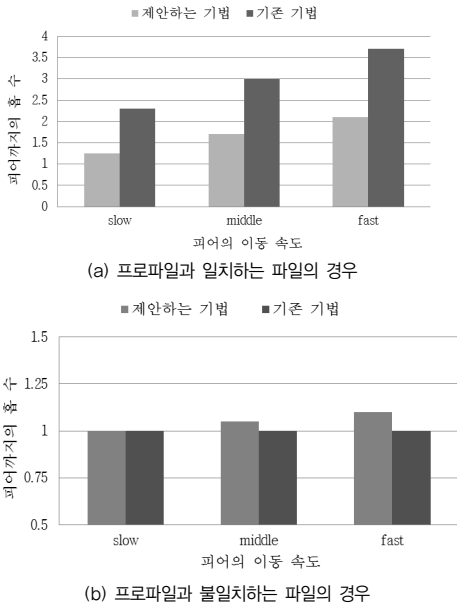


그림 6. 피어의 이동 속도에 따른 콘텐츠 검색 성능

프로파일 일치 여부만으로 데이터를 저장하던 기존 기법의 문제점을 해결하기 위해 제안하는 기법에서는 다양한 요소를 고려하여 데이터를 저장하였다. 그에 따라 프로파일과 일치하는 콘텐츠와 일치하지 않는 콘텐츠 모두의 성능 향상을 시킬 수 있다. [그림 6]은 피어의 수가 50개 일 때 피어의 이동 속도에 따른 콘텐츠 검색 성능을 나타낸다. [그림 6]의 (a)는 콘텐츠를 검색하는 피어의 프로파일과 일치하는 콘텐츠를 검색하는 경우이고 [그림 6]의 (b)는 콘텐츠를 검색하는 피어의 프로파일과 일치하지 않는 콘텐츠를 검색한 성능 평가 결과를 나타낸다. 검색 성능은 특정 콘텐츠에 대한 정보를 가지고 있는 피어까지의 홉 수를 통해서 평가하였다. 이때, 로컬에서 검색될 경우는 1-홉으로 측정하였다. 제안하는 기법과 기존 기법 모두 자신의 프로파일과 일치하는 콘텐츠에 대한 검색은 거의 1홉 내에서 검색 가능하다. 즉, 데이터 배포 과정을 통해 로컬에 저장한 인덱스들로 처리가 가능하다. 그러나 프로파일과 일치하지 않는 콘텐츠에 대한 검색의 경우에는 제안하는 기법이 기존 기법에 비해 약 40% 향상된 성능을 보이는 것을 확인할 수 있다. 이것은 제안하는 기법의 경우 자신의 프로파일과 일치하는 콘텐츠에 대한 인덱스뿐만 아

니라, 데이터의 우선 순위를 결정하는 기준에 따라 대중성을 고려해 프로파일과 일치하지 않는 콘텐츠에 대한 인덱스도 저장하고 있기 때문이다.

V. 결론

본 논문에서는 모바일 P2P 네트워크에서 피어들의 제한된 메모리, 통신 대역폭, 컴퓨팅 능력과 피어들의 이동성으로 인한 빈번한 토폴로지 변화 등의 제약점을 고려하면서도 사용자의 콘텐츠 검색 성능을 향상시킬 수 있는 데이터 배포 기법을 제안하였다. 제안하는 기법에서는 타임스탬프 메시지를 사용해 중복적인 데이터 전달을 방지하였고 제한된 메모리에 데이터를 저장하기 위한 우선 순위 결정 기법을 제안하였다. 우선 순위 결정에서는 프로파일과 일치 여부뿐만 아니라 피어 간의 연결성, 배포 정도를 고려함으로써 프로파일과 일치하는 콘텐츠뿐만 아니라 일치하지 않는 콘텐츠 검색 성능도 향상 될 수 있도록 하였다. 배포되는 데이터 크기와 검색 성능을 비교한 결과 제안하는 기법은 배포되는 데이터 크기에 비해 검색 성능이 우수한 것을 확인할 수 있었다.

향후 연구로는 브로트캐스트로 인해 발생하는 메시지를 감소시키기 위한 기법에 대한 연구를 진행할 것이다.

참고 문헌

- [1] M. Li, E. Chen, and Phillip C. Sheu, "A Chord-Based Novel Mobile Peer-to-Peer File Sharing Protocol," Proc. Asia-Pacific Web Conference, pp.806-811, 2006.
- [2] D. T. Ahmed and S. Shirmohammadi, "Design Issues of Peer-to-Peer Systems for Wireless Ad Hoc Networks," Proc. International Conference on Networking, p.26, 2007.
- [3] P. Tiago, N. Kotilainen, H. Kokkinen, M. Vapa, and J. K. Nurminen, "Mobile Search - Social

- Network Search Using Mobile Devices,” Proc. IEEE Consumer Communications and Networking Conference, pp.1201-1205, 2008.
- [4] R. Haw, C. S. Hong, and D. S. Kim, “Group P2P Network Organization in Mobile Ad-Hoc Network,” Proc. Asia-Pacific Network Operations and Management Symposium, pp.477-480, 2009.
- [5] N. Shah and D. Qian, “An Efficient Overlay for Unstructured P2P File Sharing over MANET using Underlying Cluster-based Routing,” KSII Transactions on Internet and Information Systems, Vol.4, No.5, pp.691-988, 2010.
- [6] Z. Chen, H. T. Shen, Q. Xu, and X. Zhou, “Instant Advertising in Mobile Peer-to-Peer Networks,” Proc. International Conference on Data Engineering, pp.736-747, 2009.
- [7] K. F. Yeung, Y. Yang, and D. Ndzi, “Context-Aware News Recommender in Mobile Hybrid P2P Network,” Proc. International Conference on Computational Intelligence, Communication Systems and Networks, pp.54-59, 2010.
- [8] O. Wolfson, B. Xu, and R. M. Tanner, “Mobile Peer-to-Peer Data Dissemination with Resource Constraints,” Proc. International Conference on Mobile Data Management, pp.16-23, 2007.
- [9] K. S. Bok, D. W. Kwak, and J. S. Yoo, “A Resource Discovery with Data Dissemination over Unstructured Mobile P2P Networks,” KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS, Vol.6, No.3, pp.815-834, 2012.
- [10] 광동원, 복경수, 강태호, 여명호, 유재수, 조기형, “콘텐츠 검색 서비스를 위한 효율적인 이동 P2P 구조,” 한국콘텐츠학회논문지, 제9권, 제1호, pp.30-44, 2009.
- [11] 광동원, 복경수, 박용훈, 정근수, 최길성, 유재수, “이동 P2P 환경에서 효율적인 데이터 전송을 이용한 피어 색인 기법,” 한국콘텐츠학회논문지, 제10권, 제9호, pp.26-35, 2010.
- [12] A. Klemm, C. Lindemann, and O. P. Waldhorst, “A Special-Purpose Peer-to-Peer File Sharing System for mobile Ad hoc Networks,” Proc. Vehicular Technology Conference, Vol.4, pp.2758-2763, 2003.
- [13] B. Xu, F. Vafaei, and O. Wolfson, “In-Network Query Processing in Mobile P2P Database,” Proc. ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, pp.207-226, 2009.
- [14] B. Qureshi, G. Min, D. D. Kouvasos, and M. Ilyas, “An Adaptive Content Sharing protocol for P2P Mobile Social networks,” Proc. IEEE International Conference on Advanced Information Networking and Applications Workshops, pp.413-418, 2010.
- [15] T. Brinkhoff, “A Framework for Generating Network-Based Moving Objects,” GeoInformatica, Vol.6, No.2, pp.153-180, 2002.

저 자 소개

복 경 수(Kyoung-Soo Bok)

정희원



- 1998년 2월 : 충북대학교 수학과 (이학사)
- 2000년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2005년 2월 : 충북대학교 정보통신공학과(공학박사)

- 2005년 3월 ~ 2008년 2월 : 한국과학기술원 전산학과 Postdoc
- 2008년 3월 ~ 2011년 2월 : (주)가인정보기술 연구소
- 2011년 3월 ~ 현재 : 충북대학교 정보통신공학과 초빙부교수

<관심분야> : 데이터베이스 시스템, 자료저장 시스템,

위치기반서비스, 모바일 P2P 네트워크, 센서네트워크 및 RFID 등

조 미 림(Mi-Rim Cho)

정회원



- 2010년 2월 : 충북대학교 정보통신공학과(공학사)
- 2012년 2월 : 충북대학교 정보통신공학과(공학석사)

<관심분야> : 데이터베이스 시스템, 무선 센서 네트워크, 위치기반 서비스, 클라우드 컴퓨팅 등

유 재 수(Jae-Soo Yoo)

종신회원



- 1989년 2월 : 전북대학교 컴퓨터공학과(공학사)
- 1991년 2월 : 한국과학기술원 전산학과(공학석사)
- 1995년 2월 : 한국과학기술원 전산학과(공학박사)

- 1995년 3월 ~ 1996년 8월 : 목포대학교 전산통계학과 전임강사
- 1996년 8월 ~ 현재 : 충북대학교 정보통신공학과 및 충북BIT연구중심대학육성사업단 교수

<관심분야> : 데이터베이스시스템, 정보검색, 센서네트워크 및 RFID, 멀티미디어 데이터베이스, 분산객체컴퓨팅 등