

# 대규모 데이터의 효율적인 전송을 위한 물리적 정보망을 이용한 클러스터링 기법

주상욱<sup>o</sup> 이상준

송실대학교 컴퓨터학부

jupiterpen@ssu.ac.kr, sangjun@ssu.ac.kr

## Clustering Technique using Physical Network Information for Efficient Massive Data Transmission

Sangwook Joo<sup>o</sup>, Sangjun Lee

School of Computing, Soongsil University

### 요 약

웹 2.0 환경에서 인터넷 사용자가 생성하는 정보는 폭발적인 규모로 증가하고 있다. 또한 UCC 등의 사용자 참여 서비스 및 VOD, IPTV 등의 대용량 서비스가 본격화 되고 있다. 그러나 이러한 데이터 전송량 증가 속도를 네트워크 전송 설비의 증설이 따라가지 못하고 있는 실정이다. 이를 극복하기 위해 P2P 기술을 이용하고 있지만 대부분의 P2P 기술들은 실제 물리적인 네트워크 상태를 고려하지 않고 응용 계층만을 고려하기 때문에 데이터 전송의 비효율이 발생하게 된다. 게다가 이러한 비효율을 해결하기 위한 방안들은 분산형 Pure P2P 시스템이나 구조적 P2P 시스템에 대한 연구가 대부분이고 비구조적 중앙 집중형에 대한 연구는 없는 실정이다. 본 논문에서는 물리적인 네트워크 정보와 그래프 클러스터링 기법을 적용한 계층적 클러스터링 방법을 이용하여 실제 기업에서 운영하는 중앙 집중형 P2P 시스템에서 성능을 향상시킬 수 있는 기법을 제안한다. 그리고 이를 통해 기존의 기법들이 가지고 있는 과도한 메시지 교환, 고정된 랜드마크의 유지 등의 문제점을 보완하여 대규모 데이터의 효율적인 전송을 가능케 하는 실제적인 P2P 환경에 적합한 오버레이 네트워크 모듈을 구현하였다.

### 1. 서 론

웹 2.0 환경에서 인터넷 사용자가 생성하는 정보는 폭발적인 규모로 증가하고 있다. 또한 UCC 등의 사용자 참여 서비스 및 VOD, IPTV 등의 대용량 서비스가 본격화 되고 있다. 이러한 대용량의 서비스를 지원하기 위해 네트워크 업체 등은 지속적으로 설비 용량을 증설하고 있는 실정이다. 그러나 이러한 데이터 전송량 증가 속도를 네트워크 전송 설비의 증설이 따라가지 못하고 있으며, 특히 HD급 미디어스트림 서비스 등의 초대용량 서비스 지원에 어려움을 겪고 있다. 이를 극복하기 위해 Grid Delivery, P2P 기술 등 다양한 방법이 이루어지고 있으며, 특히 현재 P2P 관련 패킷은 전세계 인터넷 트래픽의 70% 이상을 차지하고 있을 만큼 많은 비중을 차지하고 있다.

그러나 기존의 P2P 네트워크는 본질적으로 가상의 네트워크를 구성하고 있기에 실제 물리적인 네트워크와 무관하게 동작하게 되어 데이터 전송의 비효율이 발생하게 된다. 다시 말해, 임의로 구성된 P2P 오버레이 네트워크 상의 인접한 두 Peer가 물리적인 네트워크 상에서는 매우 떨어져 있을 수 있다. 이러한 상황은 네트워크 자원을 과도하게 소모하게 되고, 궁극적으로 P2P 시스템의 성능을 심각하게 저해하게 된다. 이러한 물리적인 네트워크와 오버레이 네트워크의 불일치를 해결하기 위해 다양한 연구가 시행되어 왔으나, 대부분 Gnutella와 같은 순수 P2P 모델 또는 CAN, Chord와 같은 구조적 P2P에 대한 연구가 주를 이루고 있으며, 기업 환경에 적합한

중앙 집중형 P2P 모델에 대한 연구는 미비한 실정이다.

본 논문에서는 P2P 오버레이 네트워크를 물리적인 네트워크 정보를 이용하여 동적으로 구성하여 기업환경에 적합한 중앙 집중형 P2P 시스템의 성능을 향상시킬 수 있는 기법에 대해 제안하고 구현한다. 제안 기법은 P2P 시스템에 개입하는 Peer들을 물리적 인접성에 근거하여 수평적으로 그룹핑(Grouping)하는 mOverlay[1] 방식을 이용하고, 이러한 수평적인 그룹 간의 연결 정보를 기반으로 중앙의 서버는 Graph Clustering 기법을 적용하여 계층적으로 클러스터링을 수행하며, 그룹 ID를 비트열로 부여한다. 여기서 인접한 그룹 ID 값은 인접한 그룹임을 의미하며 이러한 그룹 ID를 부여하기 위해 그래프 정보에서 Minimum Spanning Tree (이하 MST) 추출 기법을 이용하여 계층적 클러스터링을 수행한다. 이를 적용하여 기존의 기법들이 가지고 있는 데이터 전송의 비효율, 과도한 메시지 교환, 고정된 랜드마크의 유지 등의 문제점을 보완하여 대규모 데이터의 효율적인 전송에 적합한 실제 파일 공유 시스템을 구현한다.

본 논문의 구성은 다음과 같다. 2장에서 효율적인 전송을 위한 관련 연구에 대해 살펴보고, 3장과 4장에서는 각각 제안 기법과 실험적 평가에 대해 설명한다. 마지막으로 5장에서 결론을 맺는다.

### 2. 관련 연구

#### 2.1 위상의 불일치 (Topology Mismatching)

일반적으로 P2P 네트워크는 기본적으로 가상 네트워크(Virtual Overlay Network)이기에 물리적인 IP 네트워크의 상태와 무관하게 데이터 교환이 발생하게 된다. 다시 말해 IP Network Topology의 상태를 반영하지 못하는 경우, 최적화된 데이터 교환이 이루어지지 못해 시스템 전체의 성능 저하가 발생하게 된다. Peer간의 거리는 Network Delay를 사용하는 경우, 임의의 연결 오버레이 네트워크를 사용할 때 전송 효율은 매순간의 가상 연결 상태에 달려 있기 때문에 최적화가 불가능하게 된다. 이에 비해 물리적인 정보를 기반으로 클러스터링 되어 있는 경우 동일 클러스터 내에서 우선적인 검색 및 전송이 발생하게 되어 최적화가 가능하게 된다. 다시 말해, 물리적인 네트워크 연결 상태를 반영한 P2P 네트워크를 구성할 수 있다면 효율적인 P2P 전송망을 구성할 수 있게 된다.

## 2.2 물리적 위치 파악 (Location-awareness)

실제 Peer 간의 물리적 연결 정보를 이용하여 클러스터링을 구성하기 위해서는 우선적으로 접속된 Peer의 지리적 및 물리적 네트워크상의 위치를 파악해야 한다. 본 절에서는 Peer의 물리적인 위치 파악을 위해 사용되고 있는 기법들에 대해 살펴본다.

### 2.2.1 IP Geolocation

P2P 네트워크에 가입하고 있는 Peer들의 물리적인 위치를 추정하는 방법 중 사전에 조사된 물리적인 연결 정보에 기반하여 현재 접속된 Peer의 위치를 추정하는 방식이 대표적이다. 이러한 Peer 위치 파악에 대한 기법으로는 GeoPing[2], GeoCluster[3] 등이 있으며, 최신의 연구는 PlanetLab[4]과 같이 Landmark정보를 이용하여 접속된 Peer의 RTT 위상을 파악하여 이를 위치정보로 가공하여 제공하여 Peer의 위치를 추정하는 방식 등이 있다. 그러나 이러한 연구 결과는 오차 범위가 수백KM에서 100KM로 네트워크 범위가 전지구적 또는 미국과 같이 대륙 규모에 적합하며, 면적이 적은 국내의 네트워크 환경에는 적합하지 않다. 또한 실제 물리적인 거리가 멀다 할지라도 동일 ISP내의 데이터 교환이 더 빠른 현실을 제대로 반영하지 못하는 단점을 가지고 있다. 또한 고정된 Landmark를 운영해야 하며, 고정된 Landmark의 위치 자체 및 개수에 민감한 문제를 가지고 있다.

DNS 정보에 기반한 GeoTrack[5], NetGeo[6], IP2Location[7], MaxMind[8] 등의 위치 확인 기법이 있다. DNS 기반 기법은 WHOIS 정보를 기반으로 위치를 추정하는 방식으로 고정 IP에 대해서는 어느 정도 정확성이 보장되나, 특히 DSL 등의 유동 IP의 위치 정보에 대해서는 매우 정확성이 떨어진다. 예를 들어 국내 KT 등의 초고속 인터넷 가입자 IP 위치 정보는 실제 위치와 무관하게 본사의 주소를 출력하게 된다.

이외 trace정보를 기반으로 물리적인 위치 관계를 확인하는 GeoTrace[9] 기법이 있다. 이 기법은 사전에 모든 IP대역에 대해 traceroute 정보를 수집한 후 이를 분류하여 P2P 네트워크에 들어온 Peer에 글로벌 ID를 부여하는 방식이다. 이 기법은 존재하는 모든 IP에 대한 traceroute 정보를 수집하는 것은 과도한 노력이 필요하

며, 유동적인 인터넷 환경을 고려할 때 적합한 방식으로 볼 수 없다. 또한 ICMP에 기반한 traceroute 정보 수집이 원천적으로 어려운 환경임을 고려하면 최적화하기가 어려운 현실이 있다.

## 3. 제안 기법

본 논문에서는 대규모 데이터 전송에 있어 효율성을 기하기 위하여 물리적 네트워크 정보와 계층적 클러스터링 방법을 이용한다.

Peer를 클러스터에 배치하기 위해서는 우선 Peer의 Locality에 대한 정보를 수집하고, 이를 기반으로 클러스터링을 수행하게 된다. 이렇게 함으로 P2P Overlay Network와 실제 물리망 사이의 괴리를 줄일 수 있으며, 궁극적으로 불필요한 트래픽 발생을 줄여 P2P 시스템 전체의 성능을 향상시킬 수 있다.

거리 기준을 이용하여 클러스터링을 수행할 때 두 가지 방법이 가능하다. 첫 번째는 동일한 지역적 위치에 기반하여 Peer들을 클러스터링하는 방법이고, 두 번째는 동일한 네트워크 상황을 가진 Peer들을 클러스터링하는 방식이다. 이 경우 RTT, hop 등이 거리의 척도로 사용될 수 있다. 본 논문에서는 두 번째 방식에 기반하여 오버레이 시스템에 가입하는 Peer들을 RTT를 기반으로 클러스터링하는 기법을 제안한다.

### 3.1 그룹핑 및 계층적 클러스터링

제안된 기법의 기본 개념은 Network Locality를 고려하는 Hybrid Structure를 가지고 있으며, Peer들의 인접성을 평가하여 그룹으로 분류한 후 그룹을 대표하는 중심 Peer(Super Peer)가 오버레이 시스템에 가입하는 Peer들의 위치를 측정하는 Landmark역할을 수행하게 한다. 또한 중심 Peer들 전체에 대해 거리 계산을 수행하지 않고, 그룹별로 점진적으로 수행하게 하여 메시지 교환을 줄이게 되며, Peer의 오버레이 네트워크 가입/탈퇴 시 그 영향이 그룹 내로 한정되게 하였다. 생성된 그룹에서 중심 Peer들 간의 거리 계산만으로 다단계 클러스터를 구성하며, 실제 질의처리에 사용하게 된다. 그룹 내의 Peer는 중심 Peer에 대한 정보를 가지고 있으며, 전체적인 물리적 접속 정보는 중앙의 서버가 중심 Peer로부터 전달받아 클러스터 생성에 사용하게 된다.

#### 3.1.1 수평적 그룹 생성

네트워크에 가입하는 Peer들을 단순히 클러스터링 하는 것은 망의 변동성을 고려할 때 과도한 메시지 교환을 발생시키게 된다. 따라서 물리적 인접성에 근거하여 Peer들을 수평적으로 그룹핑하여 과도한 메시지 교환을 줄여야 한다. 또한 Peer의 지속 시간에 비해 그룹의 지속시간이 길기 때문에 안정성이 증대되게 된다. 그룹의 정의는 다음과 같다.

그룹 : 인접성(Network Latency, Hop, RTT, etc)이 일정 수준 이내인 Peer들의 모임.

#### 3.1.2 소속 그룹 찾기 및 그룹 생성

신규로 들어온 Peer는 어떠한 그룹에 소속되든지, 아니면 자신만의 새로운 그룹을 형성하게 된다. 신규 Peer가 특정 그룹에 소속되기 위해서는 다음의 조건을 만족해야 한다.

그룹 소속 조건: 인접 그룹과 거리 위상이 소속 그룹의 거리 위상과 동일

### 3.1.3 계층적 그룹 클러스터링

수평적인 그룹 간의 연결 정보(중심 Peer간 연결 정보)를 기반으로 중앙의 서버는 Graph Clustering 기법을 적용하여 계층적으로 클러스터링을 수행하며, 그룹 ID를 비트열로 부여한다. 인접한 그룹 ID 값은 인접한 그룹임을 의미한다. 이러한 그룹 ID를 부여하기 위해 그래프 정보에서 MST 추출 기법[10]을 이용하여 계층적 클러스터링을 수행한다.

다음은 MST를 이용한 계층적 클러스터링 과정을 나타낸다.

- 그래프에서 MST를 추출한다.
- 추출된 MST에서 최장 Edge를 중심으로 분리한다.
- 분리된 sub\_tree에 대해 최장 Edge 길이가 일정 수준 이상일 때만 b의 과정을 반복한다.

[그림 1]은 최장 Edge 거리가 5 이상으로 설정된 경우, 이러한 클러스터링 과정을 통한 그룹 ID 부여를 나타내고 있다.

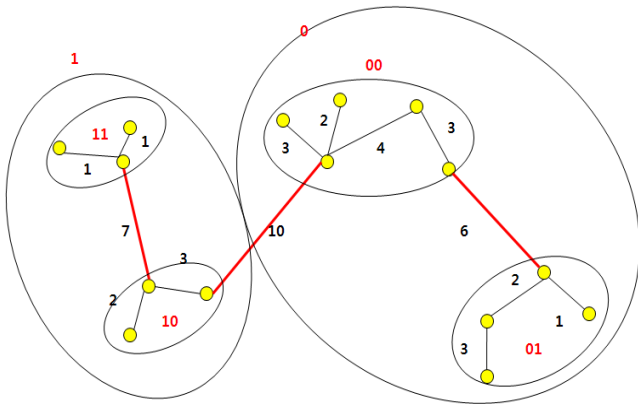


그림 1. MST를 이용한 그룹의 계층적 클러스터링 과정

MST 분할을 통해 계층형 그룹화의 결과를 그룹 ID로 부여하고, 이를 각 노드의 정보에 부여한다. 부여된 그룹 ID가 유사한 노드를 검색하고, 이를 P2P 데이터 교환에 이용한다.

부여된 그룹 ID를 기반으로 해당되는 콘텐츠를 요구하는 Peer 간의 거리계산은 XOR 연산을 통해 쉽게 구할 수 있으며, 가장 근접한 Peer로 서비스를 하도록 한다. [그림 2]는 이러한 Peer간의 거리계산을 보여주고 있다.

MST가 변경되는 경우 (새로운 그룹이 생성되거나, 또는 그룹이 제거되는 경우) 그룹 ID의 갱신이 일어나야 한다. 가장 간단한 방법은 주기적으로 재구성하는 방법이나 이는 계산량이 과도할 수 있다. 따라서 국지적으로

Group ID		Local Peer
XOR	011 001 110 011	Q
	011 001 111 001	A
-----		distance : 10
000 000 001 010		
XOR	011 001 110 011	Q
	011 001 011 101	B
-----		distance : 46
000 000 101 110		

그림 2. Peer 간의 거리 계산

이를 반영하는 기법이 필요하다. 그룹의 삭제가 발생하는 경우에는 중앙 서버에서 특별하게 할 일이 없다. 그러나 새로운 그룹이 추가되어 MST에서 Edge가 추가되는 경우에는 신규 그룹에 ID를 부여하여야 하며, 동시에 기존의 그룹 ID값들의 변경이 일어날 수 있다. 이러한 경우 신규 Edge의 길이가 MST를 분리하는 기준 값과 비교하여 국지적으로 변경을 수행한다. 즉 신규 Edge의 길이가 Cutting Boundary 이내 이면 인접 클러스터의 그룹 ID 값을 이용하며 이를 넘어서는 경우 그 클러스터에 대해서 그룹 ID 부여를 수행한다. 이러한 그룹 ID 갱신은 동적 해싱(확장 해싱) 기법을 응용하여 처리 한다.

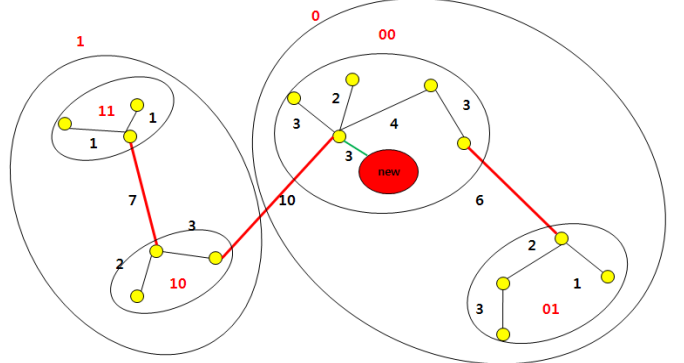


그림 3.1 신규 Edge(3) < Cluster Bound(6) 인 경우

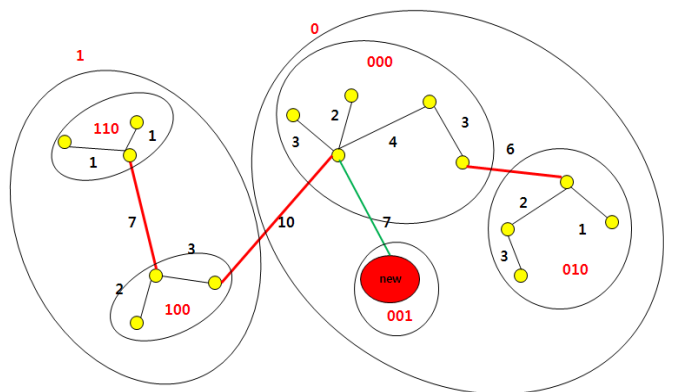


그림 3.2 신규 Edge(7) > Cluster Bound(6) 인 경우 [그림 3.1]의 경우에는 신규 Edge가 그룹 클러스터 범위 이내이기 때문에 그룹 ID의 갱신이 발생하지 않아 별도의 계산이 필요하지 않다. 그러나 문제는 [그림 3.2]

처럼 신규 Edge가 Cluster Bound를 초과하는 경우이다.

이때는 전체 그래프의 그룹 ID에 대해 변경이 일어나기 때문에 만약 전체 그래프를 재구성하게 되면 과도한 계산이 요구될 수밖에 없다. 이를 해결하기 위해 트라이트리(Trie Tree) 구조를 응용하고 0과 1로 분기를 나누어 그룹이 새롭게 생성된 곳에서만 그룹 ID를 재구성하고 그렇지 않은 범위의 클러스터는 그룹 ID에 0을 Postfix로 첨가하는 방식을 통해 계산량을 줄인다.

[그림 3]에서 변화를 살펴보면 신규 그룹 생성이 발생한 0 클러스터의 경우에는 00, 01이 각각 000, 010으로 신규 그룹은 001로 그룹 ID가 부여되었고, 신규 그룹 생성이 없는 1 클러스터의 경우에는 11, 10이 각각 110, 100으로 변경됨을 알 수 있다.

#### 4. 실험 평가

본 논문에서는 물리적 정보망과 클러스터링 기법을 응용한 오버레이 네트워크와 그렇지 않은 오버레이 네트워크의 성능 평가를 위한 시뮬레이션을 수행하였다.

실험 조건은 동일한 물리적 네트워크 크기를 갖도록 하였으며, 그 크기를 최대 50,000 노드로 설정하여 증가시켜나갔다. 물리적 네트워크는 Topology Generator인 BRITE[11]의 Graph Model을 응용하여 생성하였다. 그 외에 최대 Edge는 20으로 정하고 Clustering Boundary는 8로 설정하였으며, 노드 하나 당 취할 수 있는 Edge의 수는 최대 3개로 제한하였다. 각 노드들은 다섯 종류의 콘텐츠를 하나씩 지니며, 각각 임의의 고유IP 정보를 갖는다. 실험은 두 가지 형태로 이루어졌다. 첫 번째는 물리적 정보망을 이용한 오버레이 네트워크와 랜덤으로 생성된 오버레이 네트워크의 노드 간 거리를 측정(물리적 네트워크 크기는 50,000) 하였으며, 두 번째로는 초기 생성된 네트워크에 새로운 노드가 진입했을 시 특정 자료 검색을 위해 걸리는 시간을 측정하여 성능을 비교하였다. 실험 결과는 다음과 같다.

표 4.1 노드 간 거리 측정 결과 비교

	Random Overlay	Clustered Overlay
노드 간 거리 총합	897,644	114,185

단위 : 초

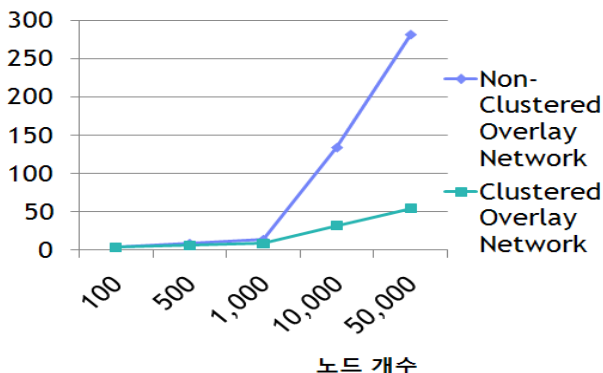


그림 4.1 특정 자료 검색 시 걸린 시간 비교

[표 4.1]의 실험 결과를 통해 물리적 정보망을 이용한 오버레이 네트워크가 그렇지 않은 네트워크 보다 월등히 실제 전송 경로를 줄일 수 있음을 알 수 있었고, [그림 4.1]의 실험 결과를 통해 클러스터링 된 오버레이 네트워크가 그렇지 않은 네트워크 보다 특정 자료 검색 시 더 우수한 성능을 나타냄을 알 수 있었다. 또한 물리적 네트워크의 크기가 급격하게 증가할수록 성능 차이가 더욱 벌어짐을 확인하였다.

#### 5. 결론 및 향후 과제

본 논문에서는 대규모 데이터를 효율적으로 전송하기 위하여 물리적 정보망과 계층적 그룹 클러스터링 기법을 사용하였다. 그리고 이를 적용하여 새로운 오버레이 네트워크 시스템을 제시하였고, 실험 평가를 통해 제안 기법의 효과를 살펴보았다.

향후 과제로 실제 네트워크 환경에서의 정확한 성능 평가와 이를 위하여 파일 공유 서비스나 미디어 스트리밍에 응용해 파일의 전송 속도를 확인하고 효율적인 트래픽 분산이 가능한 지에 대한 실험이 요구된다. 또한 제안 시스템을 데이터베이스 및 정보 검색 등의 기술과 융합하는 방법에 대한 연구도 필요하다.

#### 참고 문헌

- [1] Xin Yan Zhang, Qian Zhang, Zhensheng Zhang, Gang Song, Wenwu Zhu, "A Construction of Locality-Aware Overlay Network : mOverlay and Its Performance", IEEE Journal on selected areas in communications, vol. 22, No. 1, 2004
- [2] S. Ratnasamy, M. Handley, R.M. Karp and S. Shenker, "Topologically-Aware Overlay Construction and Server Selection", INFOCOM, 2002
- [3] T. S. Eugene Ng and Hui Zhang, "Predicting Internet Network Distance with Coordinates-Based Approaches", INFOCOM, 2002
- [4] Ethan Katz-Bassett, John P. John, Arvind Krishnamurthy, David Wetherall, Thomas Anderson, Yatin Chawathe, "Towards IP Geolocation Using Delay and Topology Measurements", SIGCOMM, 2006
- [5] Venkata N. Padmanabhan, Lakshminarayanan Subramanian, "Determining the Geographic Location of Internet Hosts", SIGMETRICS 2001
- [6] <http://www.caida.org/tools/utilities/netgeo/>
- [7] <http://www.ip2location.com/>
- [8] <http://www.maxmind.com/app/ip-location>
- [9] <http://geotrace.sourceforge.net>
- [10] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining", Addison Wesley, 2007
- [11] <http://www.cs.bu.edu/brite/>, "BRITE - Boston University Representative Internet Topology Generator."