

비상호적 비디오/오디오 분배를 위한 분산형 멀티캐스트 라우팅 알고리즘

(Distributed Multicast Routing Algorithm for Non-Interactive Video/Audio Distribution)

이 융[†] 이영석[†] 최양희^{**}

(Yung Yi) (Youngseok Lee) (Yanghee Choi)

요약 본 논문에서는 VOD(Video On Demand)나 실시간 방송 데이터 전송과 같은 비상호적 멀티미디어 응용을 위한 멀티캐스트 라우팅 알고리즘을 제안한다. 비상호적 멀티미디어 응용에서는 송신자가 다수로 존재할 수 있다. 송신자가 다수로 존재할 때 적용할 수 있는 멀티캐스트 라우팅 알고리즘을 설계하고 기존의 단 하나의 송신자가 존재할 때와 비교하여 지연 시간 및 비용 측면에서 현저한 성능향상을 가져 올 수 있음을 실험적으로 증명한다. 제안된 알고리즘의 성능 검증을 위해 광역 ATM(Asynchronous Transfer Mode)네트워크를 모델링하고, 모의 실험을 수행하였다. 모의 실험 결과 다양한 네트워크 환경 하에서 다수의 송신자 존재시 하나의 송신자만 존재할 때와 비교하여 현저한 성능향상을 보였다.

Abstract This paper proposes multicast routing algorithm that operates for non-interactive multimedia applications such as VCD(Video On Demand) and real-time data broadcasting. It is possible that there exist multiple senders in non-interactive multimedia applications. This paper designs multicast routing algorithm in this environment. The simulation of algorithm proposed in this paper shows much better performance(cost, delay) than that of only one source. To guarantee the performance of algorithm, Broadband ATM network is modelled in simulation.

1. 서론

네트워크와 컴퓨터의 빠른 발전에 기인하여 큰 대역폭을 요구하는 멀티미디어 응용 데이터의 교환이 가능하게 되었다. 대표적인 응용으로는 비디오 회의, 방송 데이터 분배, 컴퓨터 상호 협동 작업(CSCW : Computer Supported Cooperative Working), 그리고 분산 시뮬레이션 등이 있다. 이러한 응용들은 일반적으로 적은 지연 시간뿐만 아니라 지연 시간에 대한 엄격한 제한을 요구하고 있다. 또한, 네트워크의 자원을 효과적으로 이용하기 위해서 네트워크의 멀티캐스트 기능을 철저히 요구하고 있다.

ATM(Asynchronous Transfer Mode), 프레임 릴레이,

IP(Internet Protocol)와 같은 많은 패킷 스위칭 네트워크의 표준은 멀티캐스팅에 대한 지원을 포함하고 있다. 미래의 멀티미디어 응용들은 멀티캐스트 기능을 기본적으로 요구할 것이며 따라서 멀티캐스팅은 미래 네트워크의 필수적인 네트워크 구성 기술이 될 것이다. [1]

전통적으로 멀티캐스트 라우팅 문제는 하나의 송신자로부터 여러 수신자들을 연결하는 트리를 찾는 문제이다. 하지만, 본 논문에서는 같은 내용의 데이터를 송신할 수 있는 여러 개의 송신자를 둘 때 각각이 수신자를 다수의 송신자중 한 송신자와 연결하여 만들어지는 멀티캐스트 트리의 집합을 찾는 문제로 확장한다. 다시 말하면, 기존의 방법에서는 한 멀티캐스트 세션에서 생성되는 멀티캐스트 트리는 하나의 트리였으나, 본 논문에서 적용된 알고리즘을 이용하면 다수의 멀티캐스트 트리들이 생성이 되는 것이다. 본 논문에서는 응용이 요구하는 중단간 지연 시간의 제한 값을 만족하고 전체 트리의 비용을 최소화하는 트리 생성에 목적이 있다.

[†] 비 회 원 : 서울대학교 컴퓨터학부
yiyung@mmlab.snu.ac.kr
yslee@mmlab.snu.ac.kr

^{**} 종신회원 : 서울대학교 컴퓨터학부 교수
yhchoi@mmlab.snu.ac.kr

논문접수 : 2000년 8월 1일
심사완료 : 2001년 9월 7일

다수의 송신자들은 네트워크 상에서 응용 서비스 제 공시 신뢰성과 확장성을 제공한다. 즉, 기능적인 측면에서 고찰해 볼 때 다수 송신자들은 모두 동일하다고 할 수 있다. 각각의 송신자는 다수의 송신자들 중 하나의 송신자와 연결이 되어 데이터를 받게 되는데 어떤 송신자와 연결이 될지는 수신자 요구 사항 및 전체 응용의 목표에 따라 달라진다. 예를 들면 수신자는 지연 시간이 가장 짧은 송신자와 연결을 원할 수도 있고 전체 연결의 비용을 최소화하는 목표를 가지고 연결될 수도 있다.

본 논문에서는 다수의 송신자를 가진 분산형 멀티캐스트 라우팅 알고리즘을 비디오/오디오 분배에 적용하여 본다. 비디오, 오디오 분배 서비스는 VOD(Video On Demand)나 방송 데이터 분배와 같이 송신자와 수신자가 정적으로 고정되어 있고 수신자가 송신을 제어할 수 있으므로 비상호적인 응용 서비스라고 한다.

논문의 구성은 다음과 같다. 먼저 2장에서 관련 연구를 소개한다. 이어서 3장에서 본 논문에서 정의한 멀티캐스트 라우팅 문제를 소개하고, 4장에서 분산형 휴리스틱 알고리즘을 제안한다. 5장에서 IP 네트워크에서 제기한 문제해결을 위한 지원 책을 살펴보고 6장에서 시뮬레이션을 통한 성능 측정 결과를 제시한다.

2. 관련 연구

멀티캐스트 라우팅 문제는 하나의 그래프에서 한 노드(송신자)에서 다수의 노드들(수신자)까지의 연결 트리를 찾는 문제로 추상화할 수 있다. 이 멀티캐스트 트리는 송신자 기반 최단 경로 트리나 최소 비용 트리로 분류할 수 있다. 최단 경로 트리 기반 멀티캐스트 라우팅 알고리즘은 송신자로부터 수신자까지의 지연 시간을 최소로 하는 것을 목표로 하고 있으며, 다른 알고리즘에 비해서 구성하기가 매우 용이하다. 예를 들어 현재 Mbone[2]에서 사용되고 있는 DVMRP[3]는 최단 경로 기반으로 변형된 역 경로 포워딩(reverse path forwarding) 알고리즘이다.

흔히 Steiner 트리 문제로 알려져 있는 최소 비용 트리를 찾는 문제에 대한 연구도 진행되었다. Steiner 트리 문제는 NP-완전문제임이 이미 증명되었다.[4] 따라서 여러 가지 휴리스틱 알고리즘들이 트리 생성의 복잡도를 줄이기 위해서 고안되었다. 트리의 비용에 대한 연구 중 최적의 트리의 비용보다 2배를 넘지 않는 트리를 고안한 KMB 알고리즘이 대표적이다.[5]

최근에 멀티미디어 응용의 QoS를 만족시키는 알고리즘들이 고안되었다. 지연 시간 제한을 만족하면서 비용을 최소화하는 알고리즘[6], 지연 시간 제한과 더불어

지연 시간 변이의 제한을 만족시키는 알고리즘도 소개되었다[7]. 또한 각 노드의 차수 값에 대한 제한이나 대역폭에 대한 제한을 만족시키는 멀티캐스트 트리 알고리즘도 소개되었다[1][8]. 앞서 소개한 논문들은 모두 중앙 집중형 알고리즘이고 대응하는 분산형 알고리즘에 대한 연구도 발표되었다[9][10]. [11]에서 저자는 많은 QoS제한 멀티캐스트 라우팅 알고리즘의 성능을 비교 및 대조하여 어떤 알고리즘이 어떤 상황에 적합한지 분류하였다.

QoS 제한 값을 만족시키는 멀티캐스트 라우팅 알고리즘의 가능성은 자원 예약 방법에 의존적이다. 네트워크 자원을 예약하기 전에 응용에 의해서 요구되는 제한 값을 만족시키는 경로를 찾는 단계가 필요하다. 그 이후에 홉 대 홉으로 자원 예약이 진행되면 이러한 자원 예약에 대한 지원은 ATM처럼 망 자체에서 지원할 수도 있고 IP 망처럼 외부 프로토콜(RSVP)[12]을 이용하여 진행될 수도 있다.

다중 송신자나 다중 서버에 대한 연구는 최근에 보고되고 있지만, 많은 연구들이 응용 계층에서의 송신자나 서버 선택에 대한 문제를 다루고 있다[13]. 그러나, 본 논문은 네트워크 계층의 라우팅 수준에서 문제를 해결하고 있다.

3. 다중 송신자 멀티캐스트 라우팅 문제 정의

네트워크는 노드의 집합 V 와 링크의 집합 E 를 가지고 있는 유방향 그래프 G 로 형식화 할 수 있다. 집합 E 에 속한 각각의 링크들은 지연 시간과 비용의 속성을 가지고 있다. “유방향”의 의미는 링크가 비대칭이라는 의미이다. 즉, 링크 (v, w) 의 지연 시간과 비용이 링크 (w, v) 의 지연 시간과 비용과 값이 다르다는 뜻이다. 두 노드사이의 지연 시간은 보통 전과 지연 시간을 의미하고 비용은 링크의 자원 점유율의 함수이다. 자원이란 많은 것을 의미할 수 있지만, 가장 중요한 관심은 대역폭이다. 그래서 어느 한 링크의 비용을 그 링크를 점유하고 있는 트래픽의 양에 대한 함수로 계산하여도 무방할 것이다. 네트워크는 다음과 같이 모델화 할 수 있다.

- 그래프 $G=(V, E)$: 전체 네트워크
- $C_e : E \rightarrow Z^+$: 비용 가중 함수
- $D_e : E \rightarrow Z^+$: 지연 시간 가중 함수

하나의 멀티캐스트 세션에 대해서 다수의 송신자가 전체 네트워크에 분산되어 있을 경우, 수신자는 지연 시간 제한을 만족하면서 전체 트리의 비용을 최소화하도록 하나의 송신자에게 연결이 된다. 따라서, 문제를 다

음과 같이 형식화 할 수 있다.

With $G=(V, E), C_e, D_e$, 송신자 노드 그룹 $S=(s_1, s_2, s_3, \dots, s_n) \subseteq V$, 멀티캐스트 수신자 그룹 $D=(d_1, d_2, d_3, \dots, d_m) \subseteq V$, 지연 시간 제한 값 Δ

Find 최소 비용 트리의 집합, $T=(T_1, T_2, T_3, \dots, T_t), t \leq n$

Where $s \in S$ 는 같은 데이터를 송신하고 T 는 임의의 수신자 노드와 단 하나의 송신자를 연결하는데 그 수신자와 연결된 송신자 사이의 지연 시간은 Δ 보다 작다.

위의 문제는 Steiner 트리 문제로 쉽게 적용(reduction)시킬 수 있으므로 NP 완전 문제이다. 따라서 복잡도를 고려할 때 4장에서 제시할 알고리즘도 문제에 대한 최적의 해가 아닌 휴리스틱 알고리즘이다.

4. 다중 송신자 멀티캐스트 라우팅 알고리즘

4.1 중앙 집중형과 분산형 멀티캐스트 라우팅 알고리즘

멀티캐스트 트리는 두 가지 다른 형태(중앙 집중형과 분산형)로 계산될 수 있다. 중앙 집중형 알고리즘은 송신자 기반 알고리즘이라고도 불리는데, 전체 네트워크에 대한 정보를 유지하고 있는 한 중심 노드가 경로를 계산한다. 계산은 간단하고 매우 빠르다. 하지만, 중심 노드에 의한 전체 네트워크 상태 정보 유지에 대한 부하가 높고, 중심 노드에 결함이 발생했을 경우 다시 중심 노드를 선출해야 한다. 분산형 알고리즘에서는 네트워크 상태 정보의 부분적인 정보가 각 노드에서 유지되기 때문에 라우팅 테이블 유지에 대한 부하가 중앙 집중형보다 작다. 반대로 계산 시간은 중앙 집중형보다 오래 걸리게 된다. 그 이유는 경로를 계산하기 위해서 노드들 사이에서 메시지 교환이 이루어져야 하기 때문이다. 만약 트리를 계산하는 시간이 클 경우, 계산하는 동안 네트워크의 상태가 변할 확률이 높게 되고 따라서 정확한 경로 계산이 이루어지지 않을 수도 있다. 분산 알고리즘에서는 교환되는 메시지의 수와 전체 수신자를 멀티캐스트 트리에 추가하기 위한 알고리즘의 반복 회수(사이클)가 전체 계산 시간을 결정하는 두 가지 중요한 인자이므로 알고리즘은 이것을 줄일 수 있어야 한다. 시간이 중요한 인자라는 것을 고려할 때 본 논문에서 제안하는 알고리즘은 한 국면(phase)만 지나면 항상 트리를 찾을 수 있으며 교환되는 메시지의 수를 줄임으로써 계산 시간을 최소화한다.

4.2 경로 계산을 위해서 한 노드에서 필요한 네트워크 정보

각 노드에서 유지되어야 할 네트워크 정보는 {노드,

최소 지연 시간, 비용, 다음 노드}을 유지하고 있는 테이블이다. 특별히 송신자의 경우 송신자들로부터 중간 결과를 합하여 추가할 노드를 결정하는 송신자관리자에 대한 위치 정보도 유지하고 있어야 한다. 즉, 멀티캐스트 세션에 존재하는 모든 노드에 대해서 최소 지연 시간과 그 경로의 비용의 합 그리고 최소 지연 시간이 걸리는 패스의 다음 노드에 대한 정보를 유지하고 있다는 의미이다.

4.3 송신자관리자

송신자관리자는 다수의 송신자들로부터 알고리즘 실행 중간 결과를 종합하여 특정 수신 노드가 트리에 추가되도록 실제 송신 노드에게 알려 주는 역할을 수행한다. 송신자관리자의 역할에 대한 예가 그림 1에 제시되어 있다. 송신자가 3개이고 수신자가 N개일 때, M번째 수신자를 멀티캐스트 트리에 추가하는 과정이 제시되어 있다. 송신자 각각은 트리에 추가할 수신자에 대해서 최소 지연 시간과 비용을 계산하고 그 결과를 송신자관리자에게 송신한다. 송신자관리자는 그 결과를 종합하여 지연 시간 제한 값을 만족하고 비용이 가장 적게 드는 송신자가 제시한 결과를 선택하여 해당 송신자에게 선택 메시지를 보내면 그 송신자는 자신의 트리에 M번째 수신자를 추가하게 된다. 이처럼 송신자관리자는 본 논문에서 제안한 알고리즘을 실행하기 위해서 도입되는 추상적 노드이다. 본 논문에서는 송신자관리자를 선택하는 방법과 시기에 대해서는 다루지 않는다. 예를 들면 가상 탐 송신 노드는 정의된 기준에 의해서 Bully 알고리즘[14]이나 링-기반 알고리즘[15]과 같은 선거 알고리즘에 의해서 선출될 수 있을 것이다. 선택 시기는 하나의 수신 노드가 추가 될 때마다 선거 알고리즘을 실행

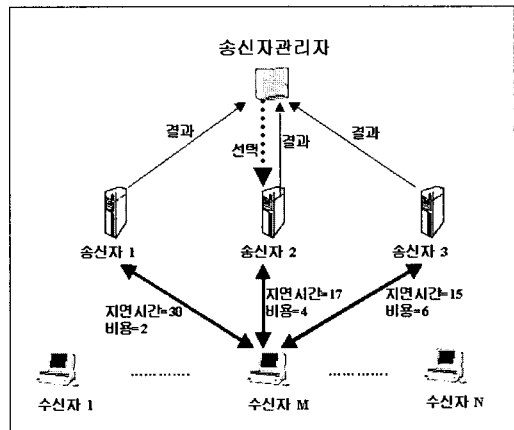


그림 1 송신자관리자의 역할

행하여 재선출 할 수도 있을 것이며, 일정 주기를 두어 그 주기에 한 번 씩 송신자관리자를 재선출 할 수도 있을 것이다.

4.4 DSGDCMT_{md} 알고리즘

본 논문에서 제안된 알고리즘을 DSGDCMT_{md}(Distributed Source Group Delay Constrained Multicast Tree - minimum delay) 이라고 명명하겠다. 본 논문은 각각의 노드들이 네트워크에 존재하는 모든 다른 노드들에 대한 최단 경로를 알고 있다고 가정한다. 각각의 멀티캐스트 수신자 및 송신자는 라우팅 테이블을 검색하면 모든 다른 노드까지의 최단 경로와 그 경로 상에서의 비용의 합을 알 수 있다. 그림 2는 DSGDCMT_{md}에서 수신자 한 노드를 멀티캐스트 트리에 추가하기 위해서 필요한 메시지 교환 절차를 보여 주고 있다.

DSGDCMT_{md} 알고리즘에서는 송신자관리자가 START 메시지를 송신자 노드들에게 보낸다. 송신자 노드들은 TEST_METRIC 메시지를 자신의 멀티캐스트 트리에 이미 포함된 노드들에게 보낸다. 트리에 아무 노드도 포함되지 않는 초기에는 그 메시지는 송신자 노드 자신에게 보내진다. TEST_METRIC 메시지를 받고 난 후, 각각의 노드들은 트리에 포함되지 않는 도착지 노드까지의 지연 시간 제한을 만족하는 패스가 있는가 살펴보기 위하여 자신의 라우팅 테이블을 검사한다. 각각의 노드들은 송신자 노드에서 자신까지의 지연시간을 알고 있고 목적지 노드까지의 최소 지연 시간과 송신자 노드-자신 지연시간을 더함으로써 송신자 노드에서 포함되어야 할 목적지 노드까지의 지연시간을 계산한다. 그 노드는 송신자-목적지 노드 사이의 지연 시간이 지연 시간의 제한 값을 만족하는가를 검사하고 그

중에서 비용이 가장 작은 경로를 선택한다. 그리고, 결과를 반영하는 정보가 TEST_METRIC_ACK 메시지에 담겨 송신자 노드에게 전달된다.

TEST_METRIC_ACK 메시지를 송신자 노드에게 보내기 전에 각각의 트리 노드들은 자신의 자식 노드들로부터 TEST_METRIC_ACK 메시지들을 종합하여 그 중 가장 작은 비용을 가지고 있는 것 하나만을 선택한다. 그리고 선택된 TEST_METRIC_ACK가 패스에서 다음 트리 노드로 전달된다. 이러한 메시지 결합(merging)은 송신자 노드로의 패스에 있는 각각의 노드에서 반복적으로 실행된다. 그러한 결과로 각각의 송신자 노드에서 수신한 TEST_METRIC_ACK 메시지들의 개수는 그 송신자 노드들이 가지고 있는 트리의 가지의 개수와 같다.

이런 과정이 모두 지난 후에 각각의 송신자 노드들은 새로운 노드를 추가시키기 위해서 어떤 경로가 선택되었는 지에 대한 정보를 가지고 있다. 각각의 송신자 노드는 그 결과를 송신자관리자에게 REQ_MERGE 메시지 형태로 보내게 된다. 송신자관리자는 모든 송신자 노드들로부터 온 그 결과들을 종합하여 그 중 가장 작은 비용 경로 결과를 가지는 송신자 노드에게 ADJUST_SRC 메시지를 보낸다.

송신자관리자에 의해서 선택된 송신자 노드는 ADJUST 메시지를 선택된 정보를 보냈던 트리의 말단 노드에게 선택되었다는 것을 알리고 그 소식은 새롭게 추가될 송신자 노드에게 ADD 메시지를 보내서 결과를 최종적으로 알린다. 새롭게 추가될 송신자 노드까지의 경로 중간에 있으면서 기존의 멀티캐스트 트리에 포함되지 않았던 노드들은 이러한 과정을 거침으로써 트리에 포함되게 된다. 이러한 노드를 본 논문에서는 중간 노드라고 한다. 중간 노드들은 ADD 메시지에 있는 축적된 지연 시간의 필드를 수정하면서 송신자 노드에서 자신까지의 지연 시간에 대한 정보를 저장한다. 그 지연 시간 정보는 자신과 그 전의 노드사이의 지연 시간을 더함으로써 실행된다. 새롭게 추가된 수신자 노드는 ADD 메시지를 받고 난 후 ADD_ACK 메시지를 보냄으로써 답한다. 그리고 선택된 송신자 노드는 ADD_ACK_SRC 메시지를 송신자관리자 노드에게 보냄으로써 마무리한다. 새롭게 수신자 노드를 추가하는 작업이 끝난 후 송신자관리자는 다시 계산될 수 있다. 계산될 수 있다는 표현은 송신자관리자를 어떤 기준에 의해서 선정하느냐에 따른다. 송신자관리자에 대한 내용은 제 4.4 절에 언급될 것이다. 본 알고리즘은 더 이상 수신자 노드가 없을 때까지 실행된다. 알고리즘의 결과는 송신

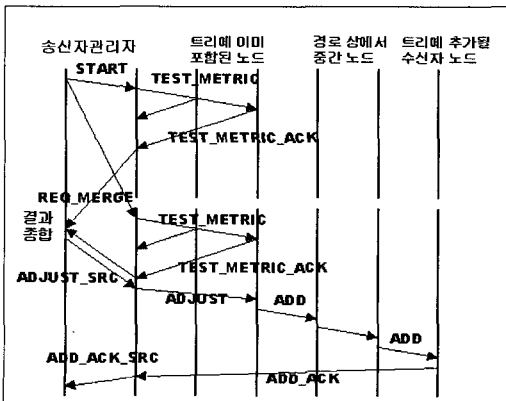


그림 2 DSGDCMT_{md}에서 메시지 교환 절차

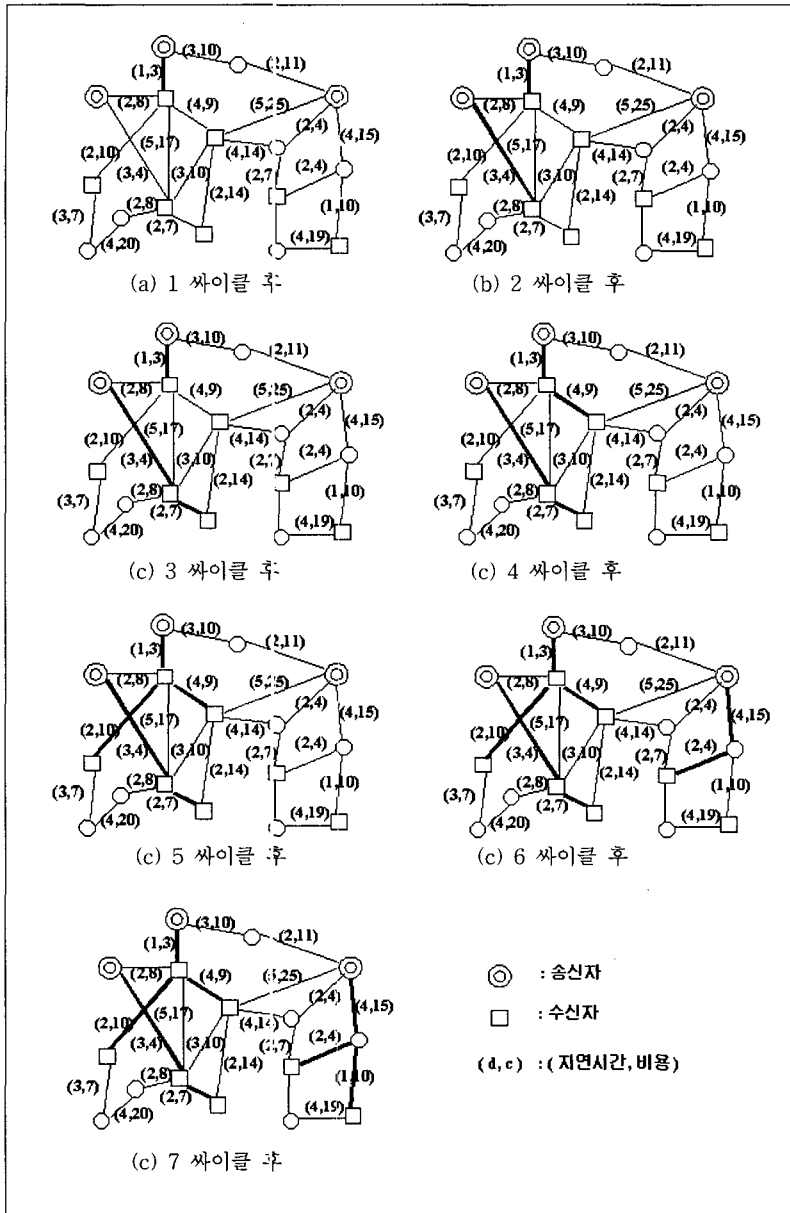


그림 3 LSGDCMT_{md}을 이용한 트리 생성의 예

자의 수와 같거나 작은 수의 멀티캐스트 트리의 집합이다. 그림 3은 DSGDCMT_{md} 알고리즘을 이용한 멀티캐스트 트리 생성과정의 예를 보여 주고 있다.

4.4.1 형식화된 알고리즘과 분석

형식화된 알고리즘 기술에 사용된 기호들은 다음과 같으며 형식화된 알고리즘은 표 1에 제시되어 있다.

- V : 네트워크 노드들의 집합
- S : 송신자 노드들의 집합
- D : 수신자 노드들의 집합, $D \subseteq V - S$
- v_T : 송신자관리자 노드
- T : 멀티캐스트 트리들의 집합

표 1 DSGDCMT_{md} 알고리즘

<p>과정 1 : v_T가 START 메시지를 $s \in S$에게 송신한다.</p> <p>과정 2 : 노드 s는 START 메시지를 받고 난 후 노드 s가 TEST_METRIC 메시지를 $v_s \in T_s$에게 송신한다.</p> <p>과정 3 : 노드 v_s는 TEST_METRIC 메시지를 수신한 다음, $w \in D - T$에 대해서 $DELAY_T(v_s) + DELAY_{\min}(v_s, w) < \Delta$를 만족하고 최소의 $COST_{\min}(v_s, w)$를 가지는 수신자 노드 w를 찾고 난 후, TEST_METRIC_ACK 메시지를 생성한다. 이 때 TEST_METRIC 메시지는 매번 한 노드를 추가할 때마다 모든 트리에 보내질 필요는 없다. 전에 수신했던 TEST_METRIC_ACK 메시지에서 새롭게 추가된 수신자 노드를 가지지 않은 트리 노드들에게 TEST_METRIC 메시지에 대한 답은 같을 것이기 때문이다.</p> <p>과정 4 : TEST_METRIC_ACK 메시지를 s에게 보내기 전에, v_s의 자식 노드들로부터 모든 TEST_METRIC_ACK들을 모아서 v_s 자신의 것도 포함하여 그 들 사이에서 최소의 $COST_{\min}(x, y)$를 가지는 한 노드를 선택한다. 이 때 $x \in T_s$와 y는 멀티캐스트 트리에 포함되지 않은 수신자 노드들이다. 선택된 메시지는 s에게 송신되고 다른 메시지들은 v_s에 의해 무시된다.</p> <p>과정 5 : 노드 s는 TEST_METRIC_ACK 메시지에서 $COST_{\min}(v_s, w)$가 최소가 되는 노드를 선택하고 REQ_MERGE 메시지를 v_T에게 송신한다. 만약, $COST_{\min}(v_s, w)$가 과정 6에서 모든 $COST_{\min}(v_i, w)$, $i \in S$에서 최소가 된다면 s와 v_s는 s_{\min}와 v_{\min}이 된다.</p> <p>과정 6 : 노드 v_T는 REQ_MERGE들을 수신한 후, 최소의 추가 비용을 가지는 $s_{\min} \in S$를 선택하고 ADJUST_SRC 메시지를 s_{\min}에게 송신한다.</p> <p>과정 7 : 노드 s_{\min}은 ADJUST_SRC를 수신한 후, ADJUST 메시지를 v_{\min}에게 송신한다.</p> <p>과정 8 : 노드 v_{\min}은 ADJUST 메시지를 수신한 후 $PATH_{\min}(v_{\min}, w)$위에 존재하는 다음 노드 x에게 ADD 메시지를 송신한다.</p> <p>과정 9 : 멀티캐스트 트리에 포함된 노드 x는 ADD를 수신한 후, $PATH_{\min}(x, w)$상에 존재하는 다음 노드 y에게 ADD 메시지를 송신한다.</p> <p>과정 10 : 노드 w는 ADD를 수신한 후, w가 트리에 포함되었다는 것을 알리기 위해 s_{\min}에게 ADD_ACK를 송신한다.</p> <p>과정 11 : 노드 s_{\min}은 ADD_ACK 메시지를 수신한 후, w가 트리에 포함되었다는 것을 알리기 위해 v_T에게 ADD_ACK_SRC를 송신한다.</p> <p>과정 12 : v_T를 재 계산한다. 하지만, 재계산은 매번 일어 날 수도 있고 그렇지 않을 수도 있다.</p> <p>과정 13 : D에 존재하는 노드가 모두 T에 포함되었으면 알고리즘은 종료되고 그렇지 않으면 과정 1에서 다시 시작한다.</p>
--

- T_i : 송신자 노드 i 의 멀티캐스트 트리에 속해 있는 노드들의 집합, $T_i \subseteq T$
- $PATH_{\min}(v, w)$: 노드 v 와 노드 w 사이의 지연 시간이 최소인 경로, $v, w \in V$
- $COST_{\min}(v, w)$: $PATH_{\min}(v, w)$ 의 비용
- $DELAY_{\min}(v, w)$: $PATH_{\min}(v, w)$ 의 지연 시간
- $PATH_{\min}(v, w, \Delta)$: 노드 v 와 노드 w 사이의 지연 시간 제한 Δ 를 만족하면서 비용이 최소인 경로, $v, w \in V$
- $COST_{\min}(v, w, \Delta)$: $PATH_{\min}(v, w, \Delta)$ 의 비용
- $DELAY_{\min}(v, w, \Delta)$: $PATH_{\min}(v, w, \Delta)$ 의 지연 시간
- $DELAY_T(v)$: 멀티캐스트 트리 T 에 존재하는 송신자 노드 s 와 노드 v 사이의 지연 시간
- Δ : 응용에 의해서 명세되는 제한된 지연 시간의 값 DSGDCMT_{md}은 지연 시간 제한 값을 만족하는 멀티캐스트 트리가 존재하면 항상 찾는다. 그리고 DSGDCMT_{md}의 메시지 교환의 복잡도는 최악의 경우 $O(|V||S|G)$ 이다.

정리 1 DSGDCMT_{md}은 지연 시간 제한을 만족하는 트리가 존재하면 항상 그 트리를 찾는다.

증명 DSGDCMT_{md}가 송신자 $s \in S$ 로부터 노드 $v_f \in D$ 까지 지연 시간 제한 경로를 찾지 못한다고 가정해 보자. 그렇다면 모든 v 에 대해서 $DELAY_T(v) + DELAY_{\min}(v, v_f) > \Delta$ 이기 때문에 $DELAY_{\min}(s, v_f) < \Delta$ 가 된다. 그런데 이 사실은 s 에서 v_f 까지 지연 시간 제한 경로가 존재한다는 가정에 모순이 된다.

정리 2 DSGDCMT_{md}알고리즘의 최악의 경우 교환되는 메시지의 복잡도는 $O(|V||S|G)$ 이다.

증명 제안된 알고리즘의 첫 단계에서 한 송신자는 트리에 이미 포함되어 있는 노드들에게 TEST_METRIC 메시지를 송신한다. 한 송신자 노드에서 한 수신자 노드까지의 링크의 수는 최대 $|V|$ 개를 가질 수 있으므로 복잡도는 $O(|V|)$ 이다. 모든 송신자가 트리에 포함되어 있는 모든 수신자에게 메시지를 보내므로 메

시지의 복잡도는 ISIG만큼 추가로 곱하여 지므로 메시지의 총 복잡도는 $O(MISIG)$ 이다.

5. 인터넷에서 다중 송신자 멀티캐스트 라우팅 알고리즘

현재 인터넷에서 적용되고 있는 IP 멀티캐스팅 구조는 많은 변형이 필요 없이 다중 송신자 멀티캐스팅을 지원할 수 있는 망이다. 다중 송신자를 지원하는 멀티캐스트 트리는 (송신자, 그룹) 형태로 인식될 수 있다. DVMRP나 MOSPF에서 처럼 각각의 (송신자, 그룹)의 조합에 분리된 트리를 사용함으로써 IP 멀티캐스팅에서 다중 송신자를 지원하게 하는 것이 가능하다. 특히, 단말 호스트는 IGMPv3[16]의 “송신자 필터링” 기능을 이용하여 특정 송신자를 선택할 수 있다. “송신자 필터링” 기능은 수신자가 데이터를 받고 싶은 송신자와 그렇지 않은 송신자를 구별하여 요청할 수 있는 기능을 의미한다.

6. 시뮬레이션을 통한 성능 평가

6.1 시뮬레이션 환경

본 논문에서는 제안된 알고리즘의 성능 평가를 위해서 모의 실험을 하였다. 모의 실험 환경은 [11]에서 만들어진 모의 실험기(이하 mcrcsim)를 참조 및 수정하여 모의 실험을 수행하였다. mcrcsim에서는 155 Mbps로 연결된 ATM 네트워크를 모델링 한다. 네트워크 노드를 배치할 배경은 미국의 전 국토를 모델로 삼아 $3000 \times 2400 \text{ Km}^2$ 의 직사각형 크기를 사용하였다. 네트워크 노드의 배치와 연결 상태는 Waxman의 그래프 생성기[5]를 사용한다. 임의로 생성된 그래프는 각각의 노드 차수를 2이상으로 되게 하였으며 전체 네트워크를 연결 그래프로 생성하였다. 블록현상이 없는 ATM 스위치를 모델링한 각 노드는 작은 버퍼를 가지고 있으면서 링크 사이의 전파 지연 시간은 빛의 속도의 3분의 2값이 되도록 한다. 송신자 노드를 위해서 VBR 송신자를 모델링 하였다. 생성된 멀티캐스트 세션이 한 링크에 보내는 트래픽의 대역폭은 임의로 할당하되 최소값과 최대값을 지정하였다. 트래픽의 대역폭은 링크의 비용 계산과 밀접한 관계를 가지고 있다. 네트워크에서 비용을 여러 가지 요소로 정의할 수 있지만, 본 논문에서는 링크의 비용에 가장 큰 영향을 미치는 요소를 링크의 대역폭과 현재 사용하고 대역폭에 대한 관계로 정의하였다. 링크의 비용 C_e 는 사용하고 있는 트래픽의 양에 비례한다. 그러므로 트래픽의 양이 많은 링크는 비용이 높고 적은

트래픽이 지나가는 링크는 비용이 낮게 된다. 기존의 모의 실험에서 비용을 임의로 할당하는 것보다는 실제 네트워크와 매우 가깝다고 말할 수 있다.

6.1.2 Waxman 그래프 생성기

그래프 생성기는 임의의 장소에 정해진 개수의 노드를 위치시키고 다음과 같은 함수를 이용하여 두 노드 사이의 확률을 다음과 같이 계산한다.

$$P_e(v, w) = \beta \exp \frac{-d(v, w)}{\alpha L}$$

$d(v, w)$: 두 노드 v, w 사이의 거리

L : 임의의 두 노드 사이의 가능한 최대 거리

α, β : 인자들, $0 < \alpha, \beta \leq 1$

어느 한 링크의 연결은 확률 P_e 에 의해서 결정된다.

α 와 β 값은 생성된 네트워크의 차수의 패턴과 연결 상태를 결정한다. α 값이 증가할수록 떨어져 있는 두 노드간의 링크가 존재할 확률이 높아지고 β 값이 증가할수록 그래프의 노드의 차수가 전체적으로 증가할 확률이 높아진다. 본 논문에서는 $\alpha=0.2$ 와 $\beta=0.65$ 를 사용하였다.

6.2 시뮬레이션 결과

모의 실험시 각각의 실험에서 무작위로 노드가 선정 되었으며 노드끼리의 연결 상태는 6.1의 확률에 따라 계산되어 연결되었다. 각각의 링크에는 무작위의 백그라운드 트래픽을 흐르게 하여 비용 값을 계산하였다. 실험의 인자는 총 네트워크의 크기, 멀티캐스트 세션의 크기(수신자의 크기), 다른 지연 시간 제한 값을 인자로 하여 진행했으며 총 네트워크의 크기나 멀티캐스트 크기일 때는 그때의 비용의 총합을 계산하였다. 지연 시간 제한 값을 변화하면서 진행된 실험에서 멀티캐스트 트리 성공도가 제한 값에 얼마만큼 의존하며 본 논문에서 제안하는 알고리즘이 멀티미디어 응용 프로그램에 적합한지를 검사하기 위하여 진행되었다. 각각의 실험은 5% 이내의 신뢰도를 유지하기 위하여 200번 실험을 해 평균값을 계산하였다.

본 절에서는 제안된 알고리즘과 비교할 대상으로 $DCMT_{md}$ 알고리즘을 선정하였다. 그 이유는 지연 시간 제한 값을 만족하는 분산 멀티캐스트 라우팅 알고리즘 중 $DCMT_{md}$ 이 가장 좋은 성능을 유지한다고 알려져 있기 때문이다[10].

그림 4는 20개의 멀티캐스트 수신자의 크기, 5개의 송신자, 그리고 200개의 전체 네트워크에 대해서 지연 시간 제한 값을 변화시키는 실험을 보여 주고 있다. 지연 제한 값이 10msec일 때 두 알고리즘은 트리 생성을 하지 못하고 있는데 이것은 지연 제한 값이 너무 작기 때문이다. $DCMT_{md}$ 알고리즘 (송신자의 개수가 1개일

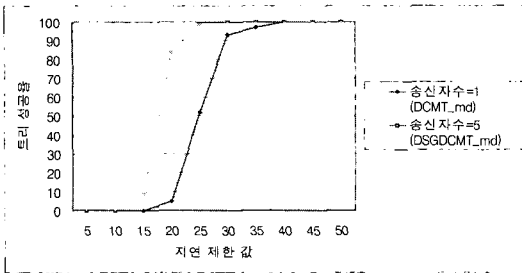


그림 4 지연 시간 제한 값의 변화에 따른 트리 성공률

때) 은 25msec일 때 약 반 정도의 성공을 거두고 있고, 40msec가 되어서야 트리 생성 성공 확률이 100%에 이르렀다. 하지만, $DSGDCMT_{md}$ (송신자의 개수가 5개일 때)는 25msec정도만 되면 트리 성공률을 100%에 가깝게 이르고 있음을 확인하고 있다. 30msec가 상호도가 높은 응용 프로그램의 지연 시간 제한 값을 고려해 볼 때 다중 송신자의 설치는 매우 중요하다고 해석할 수 있다. 특히 본 실험은 155Mbps ATM 링크를 시뮬레이션 하였기 때문에 현재 일반 망 사용자들이 접할 수 있는 망보다는 고속의 망이라고 볼 수 있다. 그러므로 본 논문에서 제안된 알고리즘은 저속 망에 적용되면 좀 더 현저한 성능 차이를 가져 올 수 있음을 짐작할 수 있다.

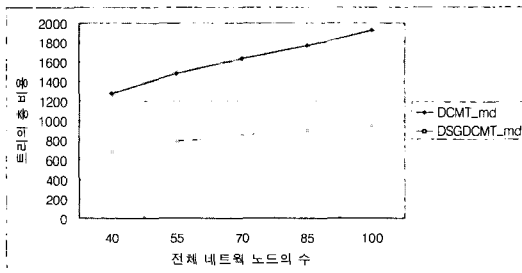


그림 5 총 네트워크 노드의 수에 변화에 따른 트리의 총 비용

그림 5는 전체 네트워크의 크기가 달라질 때 $DSGDCMT_{md}$ 와 $DCMT_{md}$ 의 트리 총 비용의 합을 보여 주고 있다. 네트워크의 크기가 작을 때($M=40$)부터 크기가 클 때($M=100$)까지 성능 향상의 정도는 대략 50%를 계속 유지하고 있다. 그림 5에서 수신자의 수는 전체 네트워크 노드의 수의 40%로 임의로 선택되었고, $DSGDCMT_{md}$ 에서 다중 송신자의 수는 5개로 고정하여 실험을 진행하였다.

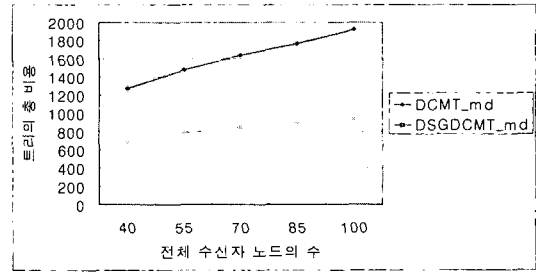


그림 6 총 수신자 노드 수의 변화에 따른 트리의 총 비용

이와 같은 성능 향상의 정도는 그림 6에서 보여 주듯이 멀티캐스트 수신자 노드 즉, 멀티캐스트 세션의 크기를 변화시킬 때도 변함없이 적용되는 것을 확인할 수 있다. 그림 6에서 전체 네트워크 노드의 수는 200개로 선정하고, 전체 송신자의 수는 10개로 고정하여 실험하였다. 좋은 라우팅 알고리즘은 확장성을 가져야 한다. 확장성을 나타내는 인자가 네트워크의 크기와 멀티캐스트 세션의 크기임을 고려해 볼 때 $DSGDCMT_{md}$ 은 특정 상황에 제한되지 않고 여러 상황에서 적용 가능함을 그림 5와 그림 6은 보여 주고 있다.

7. 결론

제한된 지연 시간을 만족시키는 멀티캐스트 트리를 생성하는 라우팅 알고리즘은 멀티미디어응용이 확산됨에 따라 그 중요성이 증가하게 되었다. 본 논문은 비상호적 응용에서 적용 가능한 다중 송신자를 고려한 멀티캐스트 라우팅 알고리즘을 제안하였다. 제안된 알고리즘은 중단간 지연 시간을 만족하면서 전체 트리의 비용을 최소화하는 트리를 찾는다. 본 논문에서 제안한 알고리즘은 분산형 알고리즘이며, 전체 네트워크의 크기와 멀티캐스트 세션의 크기에 관계없이 동일한 효율을 유지한다는 측면에서 좋은 확장성을 가지고 있다. 본 논문에서 제안한 알고리즘은 트리가 존재한다면 반드시 발견을 하고, $O(M \cdot S \cdot I \cdot G)$ 의 메시지 교환 복잡도를 가짐을 증명하였다. 또한 송신자가 하나일 때보다 좋은 성능을 보임을 모의 실험으로 확인하였다.

참고 문헌

[1] Bauer F. and Varma A., "Degree-Constrained Multicasting in Point-To-Point Networks," INFOCOM '95, 1995.
 [2] Macedonia, M.R. and D.P. Brutzman, "MBone Provides Audio and Video Across the Internet,"

- IEEE Computer Magazine, pp. 30-36, April, 1994
- [3] D. Waitzman, C. Partridge, and S. Deering, "Distance Vector Multicast Routing Protocol," RFC 1075, Nov. 1988.
- [4] P. Winter, "Steiner Problem in Networks : A Survey," Networks, vol. 17, pp. 129-167, 1987.
- [5] B. M. Waxman, "Routing of Multipoint Connections," IEEE JSAC, vol.6, no.9, pp. 1617-1622, Dec. 1988.
- [6] V. P. Kompella, J. C. Pasquale, and G. C. Polyzos, "Multicasting for Multimedia Communication," IEEE Transactions on Networkng, vol. 1, no. 3, pp. 286-292, June. 1993.
- [7] G. N. Roukas and I. Baldine, "Multicast Routing with End-to-End Delay and Delay Variation constraints," INFOCOM'96, 1996. vol. 20, no. 1, pp. 60-66, Jan. 1997.
- [8] Z. Wang, and J. Crowcroft, "Quality of Service Routing for Supporting Multimedia Applications," IEEE JSAC, vol. 14, no. 7, pp. 1228-1238, Sep. 1994.
- [9] V. P. Kompella, J. C. Pasquale, and G. C. Polyzos, "Two Distributed Algorithms for Multicasting Multimedia Information," ICC'93, 1993.
- [10] Yongjun Im, Youngseok Lee, Sunjoo Wi and Yanghee Choi, "Delay Constrained Distributed Multicast Routing Algorithm," Computer Communications, vol. 20, no. 1, pp. 60-66, Jan. 1997.
- [11] H. F. Salama, D. S. Reeves and Y. Viniotis, "Evaluation of Multicast Routing Algorithms for Real-Time Communication on High-Speed Networks," IEEE JSAC, vol. 15, no. 3, Apr. 1997.
- [12] L. Zhang, S. Deering, D. Estrin, S. Shenker and D. Zappala, "RSVP: A New Resource Reservation Protocol," IEEE Network, vol. 7, no. 5. pp. 8-18, Sep. 1993.
- [13] L. Carter, and E. Crovella, "Server Selection Using Dynamic Path Characterization in Wide-Area Networks," INFOCOM '97, 1997.
- [14] A. Silberschatz, J. Peterson, and P. Galvin, Operating Systems Concepts 4/e, Addison-Wesley, 1993.
- [15] A.S. Tanenbaum, Modern Operating Systems, Prentice Hall, 1992.
- [16] B. Cain, S. Deering, and A. Thyagarajan, Internet Group Management Protocol Version 3, Internet Draft, Feb. 1999.



이 응

1997년 2월 서울대학교 컴퓨터공학과 졸업. 1999년 2월 서울대학교 컴퓨터 공학과 석사. 1999년 2월 ~ 현재 서울대학교 컴퓨터공학부 박사과정. 관심분야는 차세대 인터넷, 무선 네트워크, 라우팅



이 영 석

1995년 2월 서울대학교 컴퓨터공학과 졸업. 1997년 2월 서울대학교 컴퓨터공학과 석사. 1997년 3월 ~ 현재 서울대학교 컴퓨터공학과 박사과정. 관심분야는 차세대 인터넷, 인터넷 트래픽 엔지니어링



최 양 희

1975년 2월 서울대학교 전자공학과 졸업. 1977년 2월 한국과학기술원 전자공학과 석사 졸업. 1984년 프랑스 E.N.S.T 대학 전산학 박사. 1981년 ~ 1984년 프랑스 CNET 연구소. 1984년 ~ 1989년 미국 IBM 왓슨 연구소, 한국전자통신연구소 책임연구원. 1991년 ~ 현재 서울대학교 컴퓨터공학과 교수