논문 2009-46TC-11-1

피어투피어 네트워크에서의 컨텐츠 기반 publish/subscribe 서비스를 위한 R-tree구성

(R-Tree Construction for The Content Based Publish/Subscribe Service in Peer-to-peer Networks)

김 용 혁*, 김 영 한*, 강 남 희**

(Yonghyuck Kim, Younghan Kim, and Namhi Kang)

요 약

컨텐츠 기반 pub/sub(publish/subscribe)서비스를 피어투피어(peer-to-peer)네트워크에서 구성할 경우 사용자의 다양한 요구 컨텐츠 정보의 분산저장 및 발생될 이벤트의 효율적인 전송이 주요 요구사항이 된다. 이와 관련하여 DHT(Distributed Hash Table)기술을 기반으로 한 pub/sub 오버레이 네트워크 구성 방식 및 사용자 요구 컨텐츠를 이용한 필터기반 트리구조의 pub/sub 네트워크 구성방식 등이 연구되어 왔다. 그러나 DHT기반 방식의 경우 토픽기반 pub/sub서비스 제공에는 적합하나 다양한 요구 조건에 따른 컨텐츠 전송에는 효율적이지 못하다. 또한 필터기반 트리구조의 방식의 경우 컨텐츠에 대한 유사 요구를 가지는 사용자가 네트워크상에 분산되어 있을 경우 비효율적인 필터구성으로 이벤트 전송 시 많은 비용이 발생한다. 본 논문에서는 다양한 사용자 요구 수용 및 이에 따른 이벤트 전송비용을 최적화하기 위해 R-Tree알고리즘 기반 오버레이 네트워크 구성방식을 제안한다. 제안된 방식은 사용자 요구의 다차원공간으로의 매핑 및 유사 요구들 간 계층적 그룹화를 통해 트리를 구성하여 비용 효율적 이벤트 전송방법을 제공하였다. 제안된 방식은 사용자요구 및 이벤트의 다양한 분포 환경에 대한 시뮬레이션을 통해 성능을 검증하였다.

Abstract

A content based pub/sub (publish/subscribe) services at the peer-to-peer network has the requirements about how to distribute contents information of subscriber and to delivery the events efficiently. For satisfying the requirements, a DHT(Distributed Hash Table) based pub/sub overlay networking and tree type topology based network construction using filter technique have been proposed. The DHT based technique is suitable for topic based pub/sub service but it's not good contents based service that has the variable requirements. And also filter based tree topology networking is not efficient at the environment where the user requirements are distributed. In this paper we propose the R-Tree algorithm based pub/sub overlay network construction method. The proposed scheme provides cost effective event delivery method by mapping user requirement to multi-dimension and hierarchical grouping of the requirements. It is verified by simulation at the variable environment of user requirements and events.

Keywords : publish/subscribe, event notify, content based routing

* 정회원, 숭실대학교 IT대학 정보통신전자공학부
(School of Information Telecommunication and Electronic Engineering, Soongsil University)
** 정회원-교신저자, 덕성여자대학교 컴퓨터공학부
(School of Information & Engineering, Duksung Women's University)
※ 본 연구는 지식경제 프론티어기술개발사업의 일환으로 추진되고 있는 지식경제부의 유비쿼터스컴퓨팅및 네트워크원천기반기술개발사업의 09C1-C2-20S과제로 지원된 것임
접수일자: 2009년8월4일, 수정완료일: 2009년11월10일

Ⅰ. 서 론

컨텐츠 기반 pub/sub(publish/subscribe)서비스는 다양한 가입자의 요구 정보를 효율적으로 전송하기위한 서비스로 다수의 가입자가 관심 이벤트에 대한 조건이 정의된 가입명세(subscription)를 시스템에 등록하고 이벤트 발생 시 이벤트에 해당되는 조건을 등록한 가입자에게만 이벤트를 전송하게 된다[1].

pub/sub 서비스 시스템은 서비스의 확장성을 고려하여 다수의 분산 서버로 구성되며 서버 간 가입명세의 분산저장 및 이에 따른 이벤트 전송을 위하여 가입명세를 이용한 오버레이 네트워크를 구성한다[2].

이와 같은 pub/sub 서비스가 p2p(peer-to-peer)네트워크에 적용될 경우 사용자의 다양한 요구 컨텐츠 정보의 분산저장 및 발생될 이벤트의 효율적인 전송이 주요 요구사항이 된다. 이는 단말 간 오버레이 네트워크를 구성하는 p2p네트워크에서의 pub/sub서비스 참여단말은 서버와 가입자 클라이언트 역할을 동시에 수행해야 함에 따라 사용자의 서비스 참여, 탈퇴는 서비스 오버레이 네트워크 토폴로지의 변동을 유발하며 이에 따른 구성 단말 간 가입자명세의 분산 또한 유동적으로 이루어져야 한다. 따라서 빈번한 네트워크 변동에 대한 빠른 대응 및 더욱 세분화된 가입명세 분산을 통한 단말 간 부하분산, 그리고 이에 따른 효율적 이벤트 전송이 요구된다[3~4].

기존 pub/sub 서비스 네트워킹 기술은 가입명세 저장 및 이벤트 전송방식에 따라 랑데부(Rendezvous) 방식과 필터방식으로 분류될 수 있다. 먼저 랑데부 방식의 경우 가입자의 가입명세 등록과 이벤트 게시자의 과 이벤트 전송이 해당 주제를 담당하는 하나의 노드를 통해 수행되는 방식인 반면 필터 방식의 경우 시스템을 구성하는 노드들이 가입명세를 분산 저장하고 가입자명세들 간의 상호 포함관계를 이용하여 노드들 간 트리를 구성한 후 이를 통해 이벤트를 전송하는 방식이다[1].

p2p환경에서 랑데부 방식의 pub/sub네트워크구성 방식에 관한 기존 연구로 DHT(distributed hash table)기반의 Ferry[5], Scribe[6], Meghdoot[7]등의 연구가 있다.

DHT방식은 p2p네트워크에서 컨텐츠 제목의 헤쉬값과 p2p네트워크 구성 노드의 ID값을 매핑 시켜 상호 연관성을 부여함으로 어떠한 중앙 제어 없이 컨텐츠의 분산 저장 및 탐색이 가능하게 한다. 이러한 DHT기반의 pub/sub네트워킹 프로토콜들은 가입명세의 헤쉬값과 매칭되는 ID를 가지는 노드로의 가입명세 등록 및 이벤트 전송을 위해서 노드 간 라우팅 방법을 제공한다. 이와 같은 방식은 모든 단말의 역할이 동일한 평면적 구조에서 토픽기반의 가입명세 즉, 가입명세가 다양한 조건을 포함하지 않고 특정 관심 토픽만 정의할 경우 가입자 명세의 분산 저장이나 이벤트 전달에 따른 부하의 분산에 있어서 좋은 특성을 가진다. 그러나 컨텐츠의 제목기반 검색에 최적화 된 DHT의 특성으로 다양한 조건으로 정의되는 가입명세 즉, 단말이 수신하고자 하는 이벤트를 정의하는 가입명세가 다양한 조건을 가지고 있는 경우 가입명세 간 연관 관계를 이용한 효율적 저장 및 이에 따른 최적의 이벤트 전달경로 탐색에 효율적이지 못하다[7].

필터기반 pub/sub네트워크 구성방법으로는 SIENA[8], Gryphon[9], REBECA[3]등의 연구가 있다. 이들 연구는 구성노드 간 연결성을 가입자명세를 이용하여 정의하며 가입명세는 네트워크상의 노드들에 분산 저장하는 방식이다. 가입명세는 네트워크상의 기존 노드 간 연결을 통해 전파되는데 전파 경로상의 노드에서 기존에 저장되어있는 가입명세가 새로운 가입명세를 포함할 경우 이를 필터링하여 더 이상 전파하지 않음으로 네트워크상의 가입명세 전파로 인한 부하를 감소시키는 동시에 효율적인 가입명세의 분산저장이 이루어진다. 가입명세가 전파된 경로는 이후 이벤트 전송 시 가입자에 대한 라우팅 경로로 사용된다. 이와 같은 방식은 가입명세가 다양한 조건으로 정의되는 경우에도 각 조건에 대한 필터 구성으로 효율적 네트워크 구성이 가능하며 유사한 가입명세를 가지는 노드들의 위치가 근접해 있을 경우 가입명세의 필터링효과가 높아져 효율적 가입명세분배 및 이벤트 전송이 가능하다. 그러나 일반적으로 노드의 위치와 가입명세는 상호 관계성이 없음으로 유사 가입명세의 노드가 분산되어 있는 일반적 상황에서 가입명세 간 필터링이 비효율적으로 이루어짐에 따라 전반적 효율이 저하된다.

이에 본 논문에서는 상기 기존 방식들의 컨텐츠 기반 가입명세에 대한 효율적 분산저장 및 이벤트 전송 문제를 해결하기 위하여 R-Tree알고리즘을 기반으로 한 pub/sub네트워크 구성방식을 제안한다. 제안되는 방식은 다수의 조건을 포함하는 가입명세를 다차원공간으로 매핑시켜 공간상의 유사 가입명세들을 R-Tree알고리즘을 이용하여 계층적으로 그룹화한 후 이를 이용하여 노드 간 트리형태의 네트워크를 구성한다. 또한 가입명세의 특성을 반영하여 노드들의 트리상에서의 역할을 정의함으로 이벤트 전송에 따른 비용을 최소화 하는

R-tree를 구성하는 방법을 제안한다.

이후 Ⅱ장에서는 이벤트, 가입명세에 대한 표현 형식이 정의되며 이를 기반으로 R-Tree구성을 위한 기본 구조, 효율적 트리 구성방안 및 이벤트 전달 절차를 기술한다. 이어 Ⅲ장에서는 시뮬레이션을 통해 R-Tree구성을 위한 다양한 방법 및 다양한 환경에 대한 특성이 비교 분석되며 Ⅳ장에서 결론은 맺는다.

Ⅱ. Publish/Subscribe R-Tree

1. 이벤트, 가입명세정의

컨텐츠 기반 pub/sub시스템에서는 가능 컨텐츠의 특성에 관한 표현방법 및 가입자가 수신하고자 하는 이벤트에 대한 정의인 가입명세 표현 방법, 게시자의 이벤트 발생 시 표현 방법이 정의되어야 한다. 본 시스템에 적용되는 pub/sub 데이터 모델 [7]에서 제안된 모델을 기반으로 다음과 같이 정의된다.

속성 구성 : $S_c = \{A_1, A_2 \cdots, A_n\}$
속성 : $A_n = \{name, \min_value, \max_value\}$
가입명세 :
$S = \{(V_{11} < A_1 < V_{12}) \wedge (A_2 = V_2) \wedge \cdots (V_n < A_n)\}$
이벤트 : $E = \{A_1 = V_1, A_2 = V_2 \cdots A_n = V_n\}$

S_c는 컨텐츠의 특성을 나타내는 속성의 집합으로 구성되며 각 속성은 속성 간 구별인자인 NAME과 해당 속성이 가질 수 있는 이 , 최대값으로 정의된다. 이와 같은 대값으 는 N개의 속성에 대의 집N차원 공간으로 모델링 되며 가입명세와 이벤트는 S_c공간 내에 영역컨텐는 단일 지점으로 위치하게 된다. 가입명세는 기본적으로 속성의 특정 관심 영역(<, >, ≥, ≤) 혹은 특정 값(=)으로 정의된다. 이와 같이 단위 속성에 대츠의 특성을는 다수의 속성과 조합되어 하나의 가입명세 최대값으게 되는데 가입명세에서 정의되지 않은 속성의 경우 해당 속성의 전 영역을 포함하는 것으로 정의된다. 반면 이벤트는 각 속성의 특정 값에 대츠의집합으로 정의된다. 이벤트 텐츠의포함하고 있지 않은 속성에 대의 포함하당 속성의 모든 영역의 값을 포함 하는 것으로 가정 모든 영와 같이 정의되는 가입명세와 이벤트는 다음 조건을 만족 할 때 상호 매칭된 것으로 가입자에게 이벤트가 NA된다.

$V_i \in A_i \ (^{\forall} i \in [1, n])$
$where\ S = \{A_1, \cdots A_n\}, E = \{V_1, \cdots V_n\}$

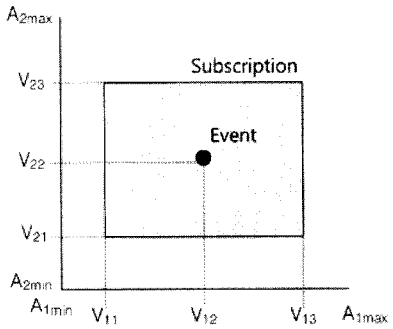


그림 1. pub/sub 가입명세 및 이벤트
Fig. 1. pub/sub subscribe and event.

그림 1은 A1 = {A1name, A1min, A1max}, A2= {A2name, A2min, A2max} 두 개의 속성으로 구성된 2차원 Sc={A1, A2}에서 가입명세 $S = (V_{11} < A_1 < V_{12}) \wedge (V_{21} < A_2 < V_{22})$가 존재하며 이벤트 $E = \{A_1 = V_{12}, A_2 = V_{22}\}$가 발생하는 예를 보인 것 이다. 그림과 같이 이벤트 발생 시 이벤트 내 모든 속성의 값이 가입명세가 정의한 영역 내에 존재할 경우 가입명세와 매칭된 것으로 가입자에게 이벤트가 전송된다.

이러한 pub/sub의 논리적 공간에 대한 정의 및 관리는 p2p단말들에 의하여 분산되어 수행되며 S_c에 대한 정의는 단말들이 기 인지하고 있다고 가정한다. 단말은 정의된 S_c공간 내에서 가입명세를 정의하여 R-Tree상의 해당 영역에 등록하게 되고 이벤트 발생 시 R-Tree 구성에 따른 계층적 오버레이 연결성을 가진 노드 간 라우팅을 통하여 이벤트와 매칭되는 가입명세를 탐색함으로 가입자에게 이벤트를 전송하게 된다.

2. R-Tree 구성

R-Tree알고리즘은 어떤 한 시스템 내에서의 효율적 데이터 객체 탐색을 위하여 유사 객체를 최소경계사각형인 MBR(Minimum Bounding Rectangle)단위로 그룹화 하여 계층적 트리를 구성하는 방법이다[10]. 단말 간 오버레이 연결을 통하여 구성되는 p2p네트워크에서 R-Tree적용 시 트리형태의 오버레이 연결 구성에 있어서의 노드의 역할에 대한 새로운 정의가 필요하다. 또한 데이터 객체가 pub/sub서비스 컨텐츠 즉, 가입명세일 때 그 특성을 고려하여 가입명세의 효율적 그룹화 방법 및 이에 따른 트리구성방법, 구성된 트리를 통한 비용 효율적 이벤트 전달 방법 등이 고려되어야 한다.

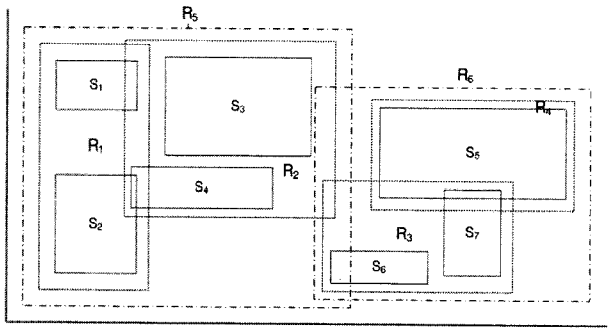
이에 본 절에서는 pub/sub R-Tree구성 시 이벤트 전송에 따른 부하를 최소화하기위한 구성 노드의 역할 정의 및 R-Tree상의 가입명세 등록 및 이벤트 전달 방

식을 기술한다.

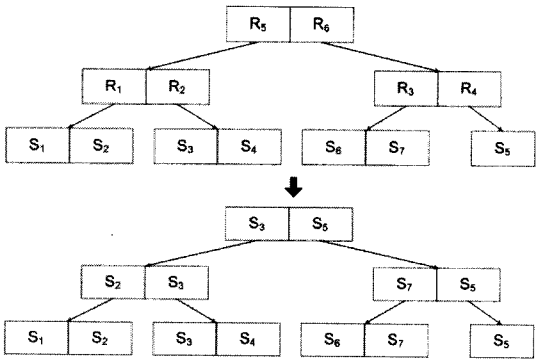
가. R-Tree 구조

S_c공간에 분포된 가입명세들은 MBR단위로 그룹화 되고 MBR간 계층적 연결을 생성함으로 트리를 구성하게 된다. R-tree알고리즘에서 트리를 구성하는 노드들은 트리 중간 노드와 최종 노드로 분류되어 최종 노드는 데이터에 대한 정보를 저장하며 중간 노드는 이러한 최종 노드 탐색에 대한 라우팅 역할만 수행한다. 그러나 가입자 노드로만 이루어진 p2p환경에서 가입명세를 가지고 있지 않고 탐색라우팅 역할만 수행하는 노드는 존재하지 않음으로 모든 노드는 가입명세를 저장하는 최종 노드 역할을 수행하는 동시에 몇몇 노드는 중간 노드 역할 또한 수행해야 한다.

그림 2.(a)는 네트워크상의 각 가입자의 가입명세 S1~S7이 분포되어 있을 때 R-Tree에 의한 트리 구성을 나타낸 것으로 R1~R4는 하위 가입명세를 그룹화한 MBR이며 R5~R6하위 MBR을 그룹화한 MBR을 나타내고 있다. (b)는 이를 기반으로 한 트리구조를 나타낸다. 그림과 같이 기존 R-tree알고리즘에 의하여 생성된 MBR기반 트리 구성에서 하위 가입명세그룹의 부모노드가 존재해야 하는데 이는 그룹화 된 가입명세중 하나가 그 역할을 수행해야 한다. 부모노드는 MBR내 다른 가입명세노드들에 라우팅을 수행한다. MBR그룹을 담당하는 부모MBR노드 또한 같은 방식으로 구성되어야 한다. 그림2.(b)의 예에서 최종계층의 S1, S2중 S2노드가 부모노드로 선출 되어 MBR R1을 관리하게 된다. 이때 대표노드 S2는 트리상의 부모노드로서 S2에 매칭되는 이벤트는 물론 S1에 해당되는 이벤트 또한 S2가 수신하여 S1로 전달한다. 이와 같이 트리상의 부모노드가 하위 모든 노드들에 대한 이벤트를 라우팅 함에 따라 적절한 부모노드 선출은 효율적 이벤트 전송 비용에 영향을 미치게 된다.



(a) 가입명세 기반 MBR구성



(b) MBR기반 트리 구성

그림 2. R-Tree 구성
Fig. 2. R-Tree construction.

MBR내 대표노드 선출방식은 다음과 같이 세 가지 방법이 있을 수 있다.

- 랜덤노드 선출
  (Random node selection : RS)
- 가장 큰 영역을 가지는 노드 서출
  (Large volume node selection: LVS)
- 오버헤드 비용이 최소화 되는 노드선출
  (Cost effective node selection : CES)

RS방식은 MBR내 노드 중 랜덤하게 대표노드를 선출하는 방식인 반면 LVS, CES방식은 가입명세 혹은 이를 그룹화한 MBR의 특성을 고려한 부모노드 선출방식이다. LVS방식의 경우 MBR내부 노드들 중 그 정의영역이 가장 큰 노드를 대표노드로 선출하는 방식이다. 그러나 CES방식은 노드의 정의영역뿐 아니라 다른 노드들의 영역과의 관계성을 고려하여 이벤트 전송에 따른 비용을 최소화하는 방식이다.

이후 절에서는 CES방식의 부모노드 선출을 위한 세부 절차 및 이를 통한 R-Tree구성방법에 대한 기술 후 상기 각 방법에 대하여 비교분석 한다.

나. R-Tree 부모노드 선출

이벤트 전송에 따른 비용은 해당 노드의 가입명세에 매칭되는 이벤트 수신처리에 대한 비용과 매칭되지 않는 이벤트의 처리 따른 비용이 있을 수 있다. 이때 매칭되지 않는 이벤트 처리는 가입자의 관심외의 이벤트에 대한 처리를 요함으로 이를 오버해드 비용으로 정의한다.

R-Tree의 이벤트 전송에 따른 MBR의 오버헤드 비용은 부모노드 가입명세와 자식노드 가입명세 간 상호

포함 관계에 따라 결정된다.

MBR내 모든 노드로의 이벤트는 트리상의 부모노드를 통하여 전송되어야 한다. 이때 부모노드의 가입명세가 자식 노드의 가입명세을 포함하고 있는 정도에 따라 부모노드의 가입명세에 관계없는 이벤트 전송률이 결정되며 이에 따른 전송 비용 즉 오버헤드 비용이 발생하며 이를 MBR의 오버헤드 비용으로 정의 할 수 있다. 이는 MBR의 부모노드를 제외한 노드들은 최종노드 역할만을 수행함으로 다른 노드들에 대한 이벤트 라우팅을 수행하지 않음에 따라 오버헤드 비용이 발생하지 않기 때문이다. 따라서 MBR영역에 해당하는 이벤트 E 전송에 있어 MBR내에서 발생하는 오버헤드 비용은 부모노드에서 유일하게 발생함으로 MBR의 전체 오버헤드 비용 $O_{MBR}(E)$는 상위노드의 오버헤드 비용 $O_P(E)$과 동일하다.

$$O_{MBR}(E) = O_P(E) \tag{1}$$

따라서 오버헤드 비용을 최소화 하는 트리를 구성하기 위해서는 MBR내부의 노드들 중 이벤트 전송에 있어서 오버헤드 비용이 최소가 되는 노드가 부모노드로 선출되어야 한다.

그림 3과 같이 S1 ~ Sn로 구성된 MBR에서 S1이 대표노드로 선출될 경우 오버헤드 비용은 인용[12]의 방법을 기반으로 다음과 같이 산출될 수 있다.

먼저 가입명세의 면적과 이벤트 수신률은 비례한다고 가정한다. 이는 가입명세가 정의하는 수신 이벤트조건의 범위가 넓을수록 범위 내에서 발생할 이벤트 확률이 높아짐으로 타당한 가정이라 할 수 있다. 또한 이벤트 수신 시 발생할 수 있는 비용은 이벤트 수신(Receive), 처리(Process), 전달(Forward)에 관한 비용만을 가정한다.

MBR내 가입명세로 이루어진 공간은 S1에만 속하는 영역(a), S1과 다른 가입명세와 겹치는 영역(b), S1이외의 영역(c)으로 분류되며 각 영역에 대한 이벤트 발생 시 이에 대 한 처리비용을 계산함으로 오버헤드 비용을 산출한다.

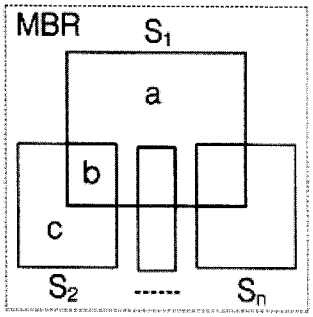


그림 3. MBR내 상위노드 선출 예
Fig. 3. The parent node selection example at MBR.

먼저 영역 a에 매칭되는 이벤트가 S1에 수신될 경우 메시지에 대한 수신, 처리비용이 발생하나 이는 S1이 정의하는 영역 내에 대한 이벤트임으로 오버헤드 비용에서 제외가 된다.

영역b에 매칭되는 이벤트 수신 시 각각 S1에 속한 영역이나 다른 가입명세에도 속한 영역임으로 수신, 처리 외에 전달비용이 추가 발생한다. 이때 수신, 처리 비용은 S1에도 매칭되는 이벤트임으로 오버헤드 비용에서 제외되고 전달비용만 오버헤드 비용으로 발생한다. 따라서 부모노드와 하위 자식노드들 간 공동영역에 대한 오버헤드 비용O_{C1}는 다음과 같이 정의된다.

$$O_{C1} = F\sum_{i=2}^{n} E(S_1 \cap S_i) \tag{2}$$

$E(\alpha)$는 α영역에 매칭되는 발생 이벤트 수를 나타낸다.

나머지 c영역에 매칭되는 이벤트 수신 시 S1에서 정의한 영역 이외의 영역에 대한 이벤트임으로 이를 수신하여 전달하는 비용이 발생하게 된다. 따라서 부모노드 영역 이외 영역에 매칭되는 이벤트 처리에 O_{E1}관한 오버헤드 비용은 다음과 같다.

$$O_{E1} = (R+F)(\sum_{i=2}^{n} E(S_i) - \sum_{i=2}^{n} E(S_1 \cap S_i)) \tag{3}$$

따라서 S1이 상위 노드로 선출 될 경우 MBR의 총 오버헤드 비용은 다음과 같다.

$$\begin{aligned} O_{MBR_1} &= O_{C1} + O_{E1} \\ &= (R+F)\sum_{i=2}^{n} E(S_i) - R\sum_{i=2}^{n} E(S_1 \cap S_i) \end{aligned} \tag{4}$$

식4로 부터 오버헤드 비용은 자식노드들의 이벤트 수신률이 낮고 공통영역에 대한 수신률이 높을 때 감소한다. 즉, 부모노드는 그 가입명세 범위가 넓으면서 하위노드와의 공통영역이 넓을 때 최소의 MBR 오버헤드 비용을 발생시킨다.

3. 가입명세 등록 및 이벤트 전달

가. 가입명세 등록 절차

(1) 가입명세 전달

가입자는 기존의 p2p 네트워크에 참여중인 노드를 통한 가입명세 등록요청 메시지 전송을 시작으로 노드 간 가입명세 라우팅을 수행하여 가입명세에서 정의하고 있는 영역을 포함하고 있는 MBR을 탐색하여 등록하게 된다. 만일 기존에 동일한 영역에 대한 가입명세가 등록되어 있을 경우 기존 가입명세의 가입자에게 새로운 가입자의 ID만 b 추가로 등록하게 된다.

Algorithm1. *Subscription _routing* (Rx_node, S_i)

Definition : Rx_Node_{MBR} *is MBR area of receiving node*
Definition : *Total_tree_level is total level of tree that is counted from root*
Require : *Subscription _routing*(n, s) *sends 가입명세 s to n node*

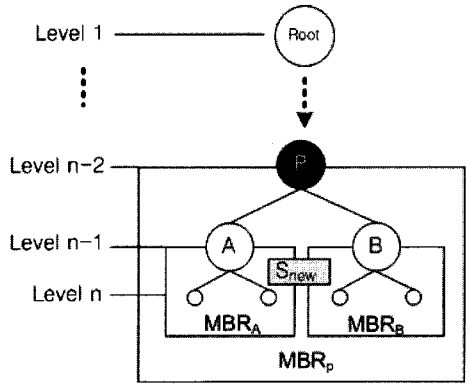
```
1: if S_i ∈ Rx_Node_MBR then
2:   if level of Rx_Node ≡ total_tree_level − 2 then
3:     Select_install_MBR(S_i)
4:   else Subscription _routing(child_node, S_i)
6:   end if
7: else Subscription _routing(parent_node, S_i)
8: endif
```

가입자로부터 가입명세 등록 요청 메시지를 수신한 기존 노드는 R-Tree상의 위치에 따라 하위 MBR을 가지는 중간노드일 경우 하위 MBR이 수신된 가입명세영역을 포함하였는지 여부를 판단하여 포함될 경우 트리의 하위 노드로 요청메시지를 전달하고 포함되지 않을 경우 상위노드로 전송하는 하게 된다. Algorithm 1은 상기 가입명세 라우팅 절차를 나타낸 것이다.

(2) 가입명세 등록 MBR선택

가입명세은 상기 라우팅 방법을 통해 최종노드로 등록되는데 이때 가입명세의 영역이 기존 MBR영역들 간 겹치는 영역에 위치할 경우 하나의 MBR을 선택하여 등록(*Select_install_MBR*)해야 한다.

기존 R-tree의 객체 삽입 방법은 객체 삽입 시 MBR 영역의 확장이 최소가 되는 MBR을 선택하여 MBR내 객체를 포함하지 않는 빈 공간을 최소화함으로 빈 공간에 대한 탐색 확률을 최소화 한다. 이는 pub/sub R-Tree에서도 동일하게 적용되는데 가입명세 등록 시 MBR의 크기가 최소로 확장되는 MBR을 선택함으로 어떠한 가입명세에도 포함되지 않는 영역에 대한 이벤트 전송에 따른 오버헤드비용 비용을 감소시킨다. 그러나 pub/sub R-Tree에서는 식4와 같이 최종 MBR을 구성하는 가입명세 간 공통 영역을 가장 많이 가지고 있



$(S_{new} \in MBR_A \,\&\, S_{new} \in MBR_B)$

그림 4. 가입명세 등록 예
Fig. 4. Subscription registration example.

는 동시에 그 영역이 가장 넓은 가입명세가 MBR의 부모노드가 될 때 오버헤드비용이 최소가 된다. 따라서 새로운 가입명세등록 시 우선적으로 MBR영역 확장이 가장 작은 MBR이 선택되고 MBR확장의 크기가 동일할 경우 가입명세가 등록되었을 때 MBR의 오버헤드 비용이 가장 작은 MBR을 선택하여 등록한다.

이와 같은 최적의 MBR선택을 위해서는 동일 계층 MBR간 오버헤드비용 비교가 이루어져야 하는데 최종 MBR의 상위 노드(total_tree_level-2) 에서 하위 MBR들로부터 오버헤드 비용을 종합하여 최적의 MBR을 선택하게 된다. total_tree_level은 루트노드(level 1)로부터 최하위 가입명세노드(level n)까지의 트리 높이를 나타낸다.

그림 4에서와 같이 MBRA, MBRB영역 모두에 포함되는 Snew가 등록되어야 할 경우 트리상의 level n-2 노드 즉, 최종 MBR(level n-1)의 부모 노드에서 하위 level n-1노드 각각에 대하여 Snew가 등록될 경우에 대한 오버헤드비용 산출을 요구한다. 이를 근거로 오버헤드 비용이 최소가 되는 MBR을 선택하여 등록 요청을 하게 된다.

각 MBR에서의 예상 오버헤드 비용은 새로운 가입명세가 등록되었을 경우를 가정 하여 산출되어야 한다. 이때 새로운 가입명세와 기존 가입명세들 간의 연관 관계로 MBR 오버헤드비용이 변화되며 이로 인하여 새로운 가입명세 또는 기존 자식노드들 중에서 새로이 부모노드로 선정되어 기존 부모노드가 변경되는 경우, 혹은 기존 대표노드변동 없이 유지되는 경우 등의 상황이 발생할 수 있다.

따라서 각 경우에 대한 MBR 오버헤드비용 산출, 즉 기존 부모노드, 새로운 가입명세 노드, 기존 자식노들이

각각 부모노드일 경우에 대한 오버헤드 비용이 산출되어야 하며 가장 작은 오버헤드비용이 MBR의 오버헤드 비용이 된다. 기존 자식노드들이 부모노드가 되었을 경우에 대한 오버헤드비용 산출 시 모든 자식노드에 대하여 산출할 필요가 없이 새로운 가입명세로부터 영향을 받는 가입명세노드들, 즉 새로운 가입명세와 공통영역을 갖는 가입명세에 대하여서만 오버헤드비용이 산출된다. 이는 새로운 가입명세와 공통영역을 갖지 않는 자식 가입명세들의 경우 식(4)에서 $(R+F)\sum_{i=2}^{n}E(S_i)$ 항에만 영향을 받는데 이항은 기존 대표노드에서도 동일한 값으로 적용됨으로 기존 대표노드보다 작은 오버헤드비용을 가질 수 없기 때문이다.

(3) 가입명세 등록

상기 절차를 통하여 선택된 MBR은 가입명세 등록 시 이전 최소 오버헤드비용 산출에 따른 부모노드선정 결과를 이용하여 기존 대표노드가 바뀌어야 할 경우 상위 부모노드와 하위 MBR내부 노드들에게 MBR 대표노드에 대한 정보를 갱신함으로 등록절차는 종료된다.

만일 가입명세에서 정의한 영역과 동일한 가입명세가 이미 존재할 경우 MBR내 가입명세의 영역을 등록할 필요 없이 기존 가입명세의 가입자에 새로운 가입자의 ID만 등록하게 된다. 이로서 동일 가입명세를 가지는 가입자 ID는 MBR내에 최초 등록된 가입자의 가입자리스트에 등록되어 이 후 가입명세에 매칭되는 이벤트 수신 시 이를 리스트 내 가입자들에게 전송한다.

나. 이벤트 전달

이벤트 전달과정은 이벤트와 매칭되는 가입명세를 탐색하여 해당 가입자에게 이벤트를 전송하게 되는 절차로 새로운 가입명세등록과정 시 가입명세전달 과정과 동일하다. 즉 이벤트 수신노드는 자신의 MBR영역에 이벤트가 매칭 되는지 여부를 판단하여 매칭 될 경우 이를 자식 노드로 전달하고 매칭되지 않을 경우 상위 부모노드로 전달하는 방식으로 이벤트를 전달하여 최종적으로 이벤트에 매칭되는 가입명세를 탐색하여 해당가입자에게 이벤트를 전송한다.

이와 같이 기본적 이벤트 전달과정은 가입명세전달 과정과 동일하나 이벤트가 MBR들 간의 공통영역에 매칭 될 경우 복수의 경로로 이벤트가 전달되는 차이점이 있다. 즉, 가입명세가 MBR간의 공통영역에 위치할 경우 적절한 MBR을 선택하여 트리의 하나의 경로를 통해 전달되는 반면 같은 경우에서 이벤트는 공통영역을 갖는 모든 MBR내의 이벤트와 매칭되는 가입명세에게 전달되어야 함으로 복수의 경로로 전달되어야 한다.

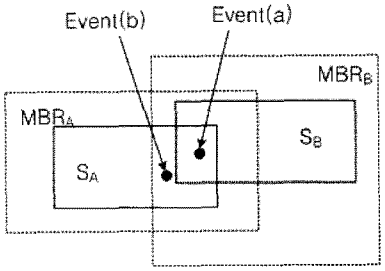


그림 5. MBR공통영역에 대한 이벤트 전달 예
Fig. 5. Event delivery example at the MBR commone area.

그림5의 예에서 이벤트 a, b 두 경우 모두 MBRA, MBRB의 공통영역에 매칭이 되어 각 MBR로 전송된다. 이벤트 a의 경우 각 MBR내 가입명세 SA, SB에 모두 매칭됨으로 가입자 A, B모두에게 이벤트가 전송된다. 반면 이벤트b의 경우 MBRA내 가입명세 SA에 매칭되나 MBRB영역의 가입명세 SB에는 매칭되지 않는다. 따라서 MBRB로 전송되었던 이벤트는 순수 오버헤드 비용으로 반영된다.

Ⅲ. 실 험

본 장에서는 가입명세와 이벤트에 대한 시뮬레이션 모델정의를 기반으로 RS, LVS, CES방식에 따른 각 pub/sub R-Tree구성 시 가입명세분포 및 이벤트 분포에 따른 오버헤드 비용, 노드에 대한 부하를 분석하였다. 또한 동일 이벤트전송의 최적화를 위한 분기노드 캐쉬 방법 적용 시 오버헤드 비용을 분석하였다.

1. 가입명세, 이벤트 모델

시뮬레이션의 Sc 및 Sc 내 각 속성은 다음과 같이 A, B 두 개의 속성으로 이루어진 2차원 공간으로 구성한다.

$$S_c=\{A_1,A_2\},\ A_1=\{A,0,1000\},\ A_2=\{B,0,1000\}$$

가입명세는 각 속성을 모두 포함하여 공간 안에 정규분포(normal distribution)에 따라 영역 및 위치하게 된다. 즉, 각 속성의 값을 정규분포에 따라 선정한 후 이들 값을 좌표로 하는 공간상의 위치가 가입명세의 중심좌표가 되며 각 속성에 대한 범위(길이)가 정규분포에 따라 선정되어 중심 좌표를 기준하여 영역을 구성한다.

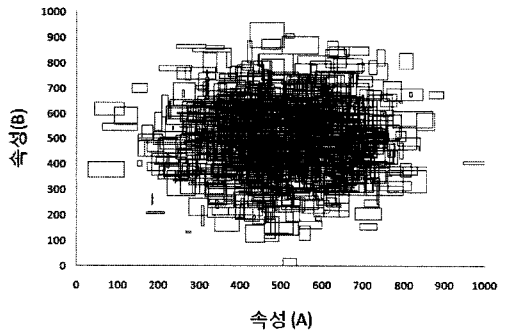
이벤트는 각 속성의 정규분포에 따른 값을 좌표로 그 위치를 지정한다.

표 1은 시뮬레이션의 가입명세 및 이벤트의 정규 분포에 대한 파라미터를 나타낸 것으로 각 파라미터는 각 속성에 동일하게 적용된다.

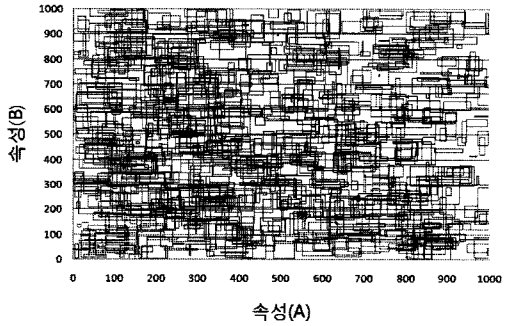
표 1에서 정의된 바와 같이 가입명세와 이벤트의 분포는 평균(m)에 의해 공간의 중심좌표(500,500)을 중심으로 분산(σ)에 따라 분포하게 된다. 그림6은 σ가 100, 300인 경우에서의 가입명세 분포를 나타낸 것이다.

표 1. 시뮬레이션 파라미터
Table 1. Simulation parameter.

Parmeter		*Value*
가입명세 위치	분포	정규분포
	평균(m)	500
	분산(σ)	100, 500
	최대, 최소	0, 1000
가입명세 영역 (속성 별 영역)	분포	정규분포
	평균	100(10%), 50(20%), 30(30%), 10(40%)
	분산	10
	최대, 최소	0, 120
가입명세 수		3000
Event	분포	정규분포
	평균	500
	분산	10 ~ 500
	최대 ,최소	0, 1000
Event 수		3000, (1000~40,000)
각 계층 (최대, 최소) 노드 수		(3, 6)



(a) σ = 100



(b) σ = 300

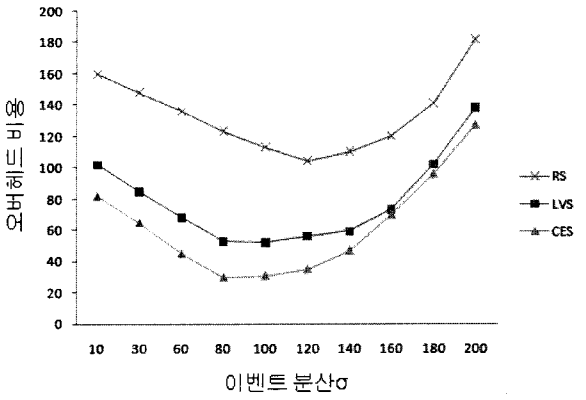
그림 6. 시뮬레이션 가입명세 분포
Fig. 6. Simulation subscription distribution.

이벤트는 R-Tree를 구성하는 노드들에서 동일한 수로 발생시키며 노드들의 가입명세와 발생시키는 이벤트와의 관계성은 없다. 이와 같은 이벤트 발생은 시뮬레이션 절차상에서 노드들의 가입명세등록절차로 R-Tree가 구성된 후 이루어지며 이벤트 발생시키는 노드의 순서는 랜덤하게 선택된다.

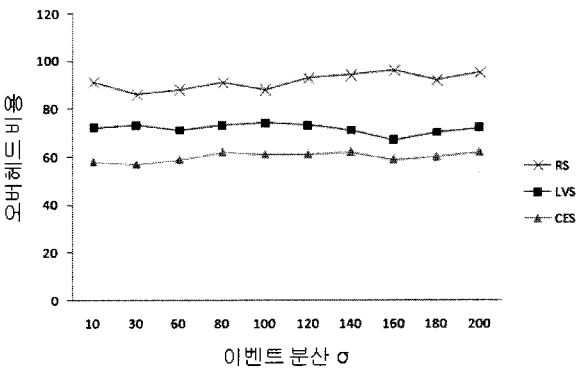
2. 가입명세, 이벤트 분산에 따른 오버헤드 비용

그림 7는 가입명세 분산이 (a)100, (b)300일 경우 각각에 대하여 이벤트의 분산의 변화에 따른 전체 노드의 평균 오버헤드 비용을 나타낸 것이다.

(a)의 경우 대체적으로 가입명세가 중앙에 집중되어 있는 상황에서 RS, LVS, CES모두 이벤트의 분산이 커질수록 오버헤드 비용이 감소하다가 다시 증가함을 볼 수 있다. 초기 이벤트의 집중된 분산(낮은 σ값)에서 높은 오버헤드 비용이 발생하는 것은 이벤트 분산이 집중



(a)가입명세 분산 σ = 100



(b) 가입명세 분산 σ = 300

그림 7. 가입명세와 이벤트 분산에 따른 오버헤드비용
Fig. 7. The overhead costs depending on distribution of the subscriptions and the events.

될수록 중앙에 밀집되어있는 다수의 가입명세에 대하여 매칭 될 확률이 높기 때문이다. 다수의 가입명세에 매칭되는 이벤트는 다수의 경로를 통하여 전송되어야 함으로 오버헤드비용이 증가하게 된다. 반대로 이벤트 분산이 넓어질수록 가입명세밀도가 낮은 영역에 매칭되는 이벤트의 개수가 늘어남에 따라 오버헤드 비용이 감소하게 된다. 이렇게 이벤트 분산이 높아질수록 오버헤드 비용은 감소하다가 σ = 120근처에서 다시 증가함을 볼 수 있다. 이는 이벤트가 가입명세분산보다 더 넓게 분산될 경우 매칭되는 가입명세가 존재하지 않는 이벤트 발생이 증가하기 때문이다. 이와 같은 현상은 R-Tree가 p2p네트워크를 기반으로 함으로 분산되어 있는 게시자들에게 모든 가입명세에 대한 정보를 사전에 알리기 어려움으로 이벤트 발생 전 매칭되는 가입명세존재 여부를 판단할 수 없기 때문이다.

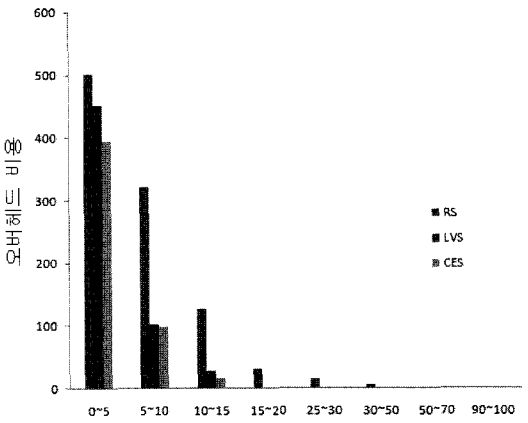
RS, LVS, CES각각의 오버헤드 비용을 비교해보면 CES가 가장 낮은 오버헤드 비용을 발생하여 RS방식과 약 두 배 정도의 차이를 보임을 볼 수 있다. 가입명세분포가 집중될 경우 가입명세영역 간 중첩이 많이 발생하게 되는데 이와 같은 가입명세간 연관관계를 이용하여 최적의 오버헤드 비용의 트리를 구성하는 CES가 RS대비 낮은 오버헤드 비용을 발생한다. LVS는 가장 넓은 영역의 가입명세가 트리상의 부모노드가 됨으로 RS와 비교하였을 때 오버헤드 비용이 낮은 가입명세가 부모노드가 될 확률이 높으나 가입명세 간 중복영역을 고려하지 않음으로 CES보다는 높은 오버헤드 비용의 부모노드가 선출될 확률이 높다. 따라서 LVS는 RS와 CES대비 중간의 오버헤드 비용을 발생시킴을 볼 수 있다.

(b)의 경우 가입명세분산이 넓게 분포되어있는 상황에서 이벤트의 분산과 관계없이 RS, LVS, CES모두 거의 일정한 오버헤드 비용을 발생시킴을 볼 수 있다. 이는 가입명세가 공간상에 균등하게 분포됨에 따라 좌표따른 가입명세의 밀집도 차이가 없음으로 이벤트의 좌표에 따른 오버헤드 비용의 차이가 없기 때문이다.

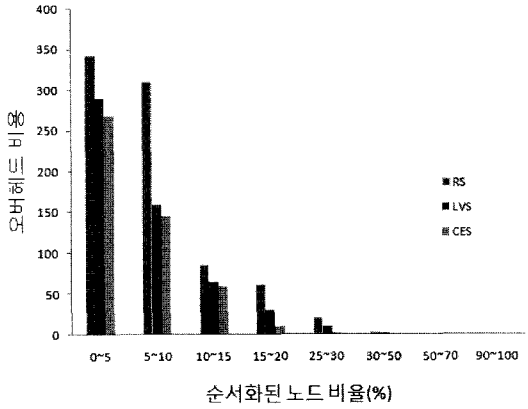
(b)의 경우에도 CES가 가장 낮은 오버헤드 비용을 발생한다. 이는 가입명세가 균등하게 분포 됨에도 각 가입명세의 넓이가 다양함으로(σ = 100~10) 중복영역이 존재하여 이를 고려한 CES가 최저 오버헤드 비용의 부모노드 선출이 가능하기 때문이다.

3. 오버헤드 비용에 대한 노드 간 부하 분산

그림 8의 실험 (a) (b)는 가입명세와 이벤트의 분산



(a) 가입명세 σ = 100, event σ = 100



(b) 가입명세 σ = 300, event σ = 300

그림 8. 가입명세 분산에 따른 노드부하
Fig. 8. The load of nodes depending on the subscription distribution.

이 각각 (100,100), (300,300)으로 분포되어있을 경우 노드 이벤트 전송 오버헤드 비용에 따른 부하 분산 정도를 실험한 것으로 노드들을 오버헤드 비용이 높은 순서로 순서화 되어 5%단위로 그룹화 한 후 각 그룹의 평균 오버헤드 비용을 나타낸 것이다.

(a)의 경우 CES가 전체노드의 15%노드에서 오버헤드 비용이 발생하며 10%노드들에 집중화 되어있음을 볼 수 있다. 반면 RS의 경우 15%정도의 노드들에 오버헤드 비용이 분산되어 있으나 CES대비 오버헤드비용순위 그룹 간 차가 크지 않다. 즉 RS의 오버헤드 비용의 분산이 CES보다 높음을 볼 수 있다. 이와 같이 CES에서의 오버헤드 비용 집중화현상은 트리의 여러 계층에 중복적으로 위치하는 노드의 높은 이벤트 전송확률로 인하여 발생한다. 즉, 이벤트 전달 시 트리상의 많은 노드들을 거치지 않고 중복된 노드를 거칠 확률이 높음으로 그 중복된 노드에 오버헤드 비용이 집중된다. 예를

들어 그림2의 (b)에서 S5와 같이 트리에서 각 계층에 중복적으로 위치하는 경우 S5가 하위 노드로 이벤트를 전송할 확률이 높음으로 이에 따른 오버헤드 비용이 S5에 집중되어 발생할 수 있다. 반면 RS의 경우 각 계층의 노드가 중복될 확률이 낮음으로 다수의 노드가 이벤트 전송에 참여하여 오버헤드 비용이 다수의 노드에 분산 발생하게 된다.

그러나 순위그룹 내 각 방식의 평균오버헤드 비용을 비교해보면 RS가 CES대비 높은 비용을 나타내고 있다. 이는 5.2절의 시험을 통해 검증된 CES의 비용 효율적 트리방식에 기인하기도 하지만 여러 계층에 중복된 노드에서 오버헤드 비용감소가 일어나기 때문이다. 여러 계층에 중복된 노드는 이벤트가 전달될 하위 계층의 노드가 자신일 경우 실 전송 없이 내부처리로 이벤트 전달이 이루어짐으로 이에 따른 전송, 수신에 대한 오버헤드 비용이 발생하지 않는다.

실험 (b)의 경우 (a)의 결과 대비 CES의 오버헤드 비용의 분산이 넓게 분포한다. 이는 가입명세와 이벤트가 넓은 영역에 분산되어 있어 중복노드를 통한 이벤트 전송이 (a)보다 적기 때문이다.

Ⅳ. 결 론

본 논문에서는 p2p네트워크에서의 pub/sub서비스를 위한 R-Tree구성방안을 제안하였다. 제안된 방안은 컨텐츠 기반 pub/sub서비스를 제공하기 위하여 다양한 속성을 통한 사용자 가입명세가 정의 되며 이를 다차원 공간에 매핑시킴으로 가입명세 영역을 구성하였다. 이를 기반으로 pub/sub서비스에 참여하는 노드들의 가입명세에서 정의 하고 있는 관심 영역을 고려하여 이벤트 전송비용에 최적화된 R-Tree트리를 구성 하였으며 시뮬레이션을 통하여 가입명세와 이벤트의 분포에 따른 오버헤드 비용이 분석되었다. 시뮬레이션 분석결과 가입명세의 분포에 따라 이벤트 전송 오버헤드 비용의 특성이 차이가 나는데 특히 가입명세의 영역 간 공통영역이 많이 발생되는 밀집된 가입명세분포에서 가장 낮은 오버헤드 비용이 발생됨을 볼 수 있었으며 이와 같은 가입명세분포는 실 pub/sub서비스 환경에 가장 근접한 형태로 제안된 방식의 실 서비스 적용에 대한 적합성을 보였다.

참 고 문 헌

[1] Liu, Y., Plale, B.: "Survey of publish/subscribe event systems," In: Indiana University Computer Science Technical Report TR-574. (2003).

[2] Carzaniga, A., Rosenblum, D., Wolf, A."Design and Evaluation of a Wide-Area Notification Service," ACM Transactions on Computer Systems 3 (Aug 2001) p.332-383

[3] Terpstra, W.W., Behnel, S., Fiege, L., Zeidler, A., Buchmann, A.P."A peer-to-peer approach to content-based publish/subscribe," In: Proceedings of the 2nd International Workshop on Distributed Event-Based Systems (DEBS'03). (2003).

[4] Costa, P., Frey, D."Publish-subscribe tree-maintenance over p2p network," In: Proceedings of the International Workshop on Distributed Event- Based Systems (ICDCS/DEBS'05). (2005).

[5] Y. Zhu and Y. Hu, "Ferry: An architecture for content-based publish/subscribe services on P2P networks," in Proceedings of the International Conference on Parallel Processing (ICPP), June 2005.

[6] M.Castro, P.Druschel "SCRIBE: A large-scale and decentralized application-level multicast infrastructure," IEEE journal on selected Areas in Communications 2002 p.100-110

[7] Gupta, A., Sahin, O., Agrawal, D., Abbadi, A.E." Meghdoot: Content-based publish:subscribe over p2p networks," In: Proceedings of the ACM/IFIP/ USENIX 5th International Middleware Conference (Middleware'04). (2004)

[8] SIENA Web Site: (http://www.cs.colorado.edu/users/carzanig/siena/)

[9] Gryphon Web Site: (http://www.research.ibm.com/gryphon/)

[10] Costa, P., Frey, D."Publish-subscribe tree-maintenance over p2p network," In: Proceedings of the International Workshop on Distributed Event- Based Systems (ICDCS/DEBS'05). (2005).

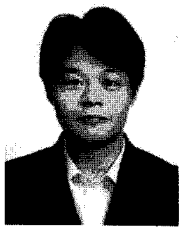
[11] Guttman, A. "R-trees a Dynamic Index Structure for spatial Searching," Proc ACM SIGMOD, 47-57, 1984.

[12] Yongqiang Huang, Hector Garcia-Molina, "Publish/Subscribe Tree Construction in Wireless Ad-Hoc Networks," LNCS 2574, pp. 122-140, 2003.

저 자 소 개

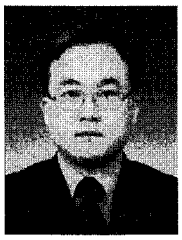
김 용 혁(정회원)

1999년 숭실대학교 정보통신공학 학사졸업
2001년 숭실대학교 정보통신공학 석사졸업
2003년 숭실대학교 정보통신공학 박사 수료
2003년~2006년 (주)코어세스 통신연구소 선임연구원
2007년~현재 숭실대학교 정보통신전자공학부 박사과정
<주관심분야: 컴퓨터 네트워크, 인터넷 네트워킹, 이동 데이터 통신망>



강 남 희(정회원)

1999년 숭실대학교 정보통신공학 학사졸업
2001년 숭실대학교 정보통신공학 석사졸업
2005년 Siegen University(독일) 컴퓨터공학 박사
2006년 다산네트웍스 선임연구원
2007년 가톨릭대학교 컴퓨터정보공학부 초빙교수
2009년 덕성여대 컴퓨터공학부 전임강사
<주관심분야 : 컴퓨터 통신, 인터넷보안>



김 영 한(정회원)

1984년 서울대학교 전자공학 학사 졸업
1986년 한국과학기술원 전기전자공학 석사 졸업
1990년 한국과학기술원 전기전자공학 박사 졸업
1987년~1994년 디지콤정보통신연구소 데이터통신연구부장
1994년~현재 숭실대학교 정보통신전자공학부 정교수
<주관심분야: 컴퓨터 네트워크, 인터넷 네트워킹, 이동 데이터 통신망>