

<http://dx.doi.org/10.7236/IIBC.2015.15.5.75>

IIBC 2015-5-9

데이터 액세스 확률의 제곱근 법칙을 이용한 상호 관련 데이터 할당 기법

An Interdependent Data Allocation Scheme Using Square Root Rule of Data Access Probability

권혁민*

Hyeokmin Kwon*

요약 데이터 할당 기술은 데이터 방송 시스템의 성능을 향상시키기 위해서 필수적이다. 본 논문은 질의 프로파일과 질의 요청 확률이 주어진 환경에서 다중 데이터 질의를 처리하기 위하여 방송채널에 데이터를 할당하는 주제를 연구하여 IDAS(Interdependent Data Allocation Scheme)로 명명된 새로운 데이터 할당 기법을 제안한다. 제안된 기법은 각 데이터의 방송빈도를 자신의 상대적 액세스 확률의 제곱근 값에 비례하게 설정하는 전략을 채택한다. IDAS 기법은 요청 확률이 높은 질의들을 빠르게 처리할 수 있고 적절한 수준의 질의 데이터 인접성을 보일 수 있기 때문에 질의 응답 시간의 성능을 향상시킬 수 있다. 제안된 기법의 성능 평가를 위해 시뮬레이션이 수행되었다. 실험 결과에 따르면, 평균 응답시간의 성능에서 IDAS는 다른 기법보다 우수한 성능을 보인다.

Abstract A data allocation technique is essential to improve the performance of data broadcast systems. This paper explores the issues for allocating data items on broadcast channels to process multiple-data queries in the environment where query profiles and query request rates are given, and proposes a new data allocation scheme named IDAS. The proposed scheme employs the strategy that the broadcast frequency of each data is determined by the square root value of its relative access probability. IDAS could enhance the performance of query response time since it can process queries of high request rate fast and show a reasonable degree of query data adjacency. Simulation is performed to evaluate the performance of the proposed scheme. The simulation results show that IDAS outperforms other schemes in terms of the average response time.

Key Words : Data allocation, broadcast schedule, simulation

1. 서론

무선 이동통신 기술과 모바일 단말 장치의 발달로 인하여 사용자가 언제 어디서나 다양한 서비스에 접근할 수 있는 모바일 컴퓨팅(mobile computing) 환경이 보편

화되고 있다. 모바일 컴퓨팅 환경에서는 상향과 하향의 비대칭적 통신 대역폭과 모바일 클라이언트의 제한된 배터리 용량으로 인하여 요청-응답(request-response) 형태로 데이터를 전달하는 방법보다는 방송 방식에 기초한 데이터 전달 방법이 훨씬 바람직하고 효율적인 데이터

*종신회원, 세명대학교 정보통신학부
접수일자 : 2015년 8월 15일, 수정완료 2015년 9월 15일
게재확정일자 : 2015년 10월 9일

Received: 15 August, 2015 / Revised: 15 September, 2015 /
Accepted: 9 October, 2015

*Corresponding Author: hmkwon@semyung.ac.kr
School of Information and Communication System, Semyung
University, Korea

전송 기술로 알려져 있다^[1]. 방송 방식의 데이터 전달 방법을 채택한 정보 시스템에서 방송 서버는 한정된 방송 대역폭을 통하여 많은 데이터들을 순차적으로 방송해야 하기 때문에 클라이언트가 어떤 데이터를 액세스하려면 해당 데이터가 방송채널에 나타나기를 기다려야 한다. 데이터 방송 시스템의 주요 성능 지수는 질의 평균 응답시간(response time)인데, 질의 응답시간이란 클라이언트가 특정 질의를 제기한 순간부터 자신이 원하는 모든 데이터들을 수신할 때까지 걸린 시간을 의미한다.

데이터 할당 기법에서는 한정된 방송 대역폭을 효율적으로 사용하기 위하여 다수의 데이터들을 어떻게 분할하여 방송채널에 할당하고, 할당된 데이터들은 어떤 빈도 및 어떤 순서로 방송할 것인가의 주제를 연구하는데, 질의 응답시간의 성능은 이 데이터 할당 기법에 큰 영향을 받는다. 이에 따라 이에 대한 많은 연구들이^[1-11] 수행되었는데, 클라이언트가 제기하는 질의의 특성에 따라 하나의 데이터를 액세스하는 단일 데이터 질의를 위한 기법들과 다수의 데이터를 액세스하는 다중 데이터 질의를 위한 기법들로 분류할 수 있다. 그리고 데이터 할당 기법은 방송채널의 수에 따라 단일 방송채널과 다중 방송채널 환경의 기법으로 분류할 수 있는데, 단일 채널 환경은 질의 응답시간의 개선에 한계가 있기 때문에 본 연구에서는 다중 방송채널 환경에 한하여 연구를 수행한다.

단일 데이터 질의 환경을 위한 대표적인 기법에는 DP(Dynamic Programming)^[2]와 RDP(Restricted Dynamic Programming)^[3] 기법들이 있는데, 이 기법들은 각 방송채널이 평형(Flat) 방송된다는 가정 하에 응답시간이 최소화되도록 데이터를 분할하여 각 채널에 할당했다. 비록 이 기법들이 방송채널별 평형 방송을 하는 시스템 모델에서는 최적 성능을 발휘하지만, 같은 채널에 할당된 데이터들을 그들의 인기도에 관계없이 동일 빈도로 방송하기 때문에 응답시간의 성능을 향상시키는 데는 한계가 있었다. TLDP(Two-Level Dynamic Programming)^[4]와 NODA(Near Optimal Data Allocation)^[5] 기법은 이 단점을 극복하기 위하여 각 방송채널은 동일 대역폭을 갖는 다수의 서브채널로 구성된다는 개념을 도입하여 서브채널 별로 평형 방식으로 방송하는 시스템 모델을 정립했다. 그리고 각 서브 채널에 데이터를 최적으로 할당하는 방법들을 개발하여 단일 데이터 질의의 이론적 최적 응답시간에 거의 필적하는 성능을 실현했다.

모바일 환경에서 개개의 데이터들은 독립적으로 액세스

스되기도 하지만, 각각의 데이터들이 서로 관련성이 있어 이들이 동시에 요청되기도 한다. 예를 들어, 교통 정보 시스템에서 교통 상황을 고려한 빠른 길을 찾기 위해서는 다수 개의 도로에 대한 교통 상황 정보를 검색하는 경우가 있고, 주식 매매 시스템에서는 의사 결정을 위하여 다수 개 업체의 주식 가격 정보를 동시에 요청하는 경우가 있을 것이다. 이와 같은 경우에 클라이언트는 다수 개의 데이터를 동시에 요청하는 다중 데이터 질의를 제기하게 될 것이다. 이와 같은 다중 데이터 질의 환경을 위해 PBA(Placement-Based Allocation)^[6]와 MA2D^[7] 기법을 비롯한 다수의 데이터 할당 기법들이^[8-9] 제안되었는데, 이들은 모두 질의가 액세스하는 데이터 집합을 의미하는 질의 프로파일과 질의 요청 확률이 제시된 환경에서 방송채널에 데이터를 할당하는 연구를 수행하였다.

PBA 기법은 질의가 필요로 하는 데이터들이 동시에 방송되는 것을 방지하기 위하여 각 방송채널의 데이터 할당 위치에 기초한 알고리즘을 개발하여, 데이터 충돌 확률을 획기적으로 감소시켜 질의 응답시간을 개선할 수 있었다. 그러나 이 기법은 한 방송주기에서 모든 데이터들을 한 번만 방송하는 정책을 사용하기 때문에 응답시간을 개선하는 데는 한계가 있었다. MA2D 기법은 이 단점을 개선하기 위하여, 질의 프로파일과 질의 요청 확률에서 각 데이터의 상대적 요청 횟수를 파악한 후 이를 정규화하여, 한 방송주기에서 각 데이터를 자신의 정규화 값에 비례하게 할당하는 기법을 제안했다. 그러나 이 기법은 요청 횟수에 너무 큰 가중치를 부여하는 경향이 있기 때문에 요청 확률이 낮은 질의의 응답시간이 너무 길어지는 단점이 있다.

본 논문에서는 이와 같은 단점을 극복하기 위하여, 각 데이터의 액세스 확률에 적절한 가중치를 부여하는 방법을 고안하여, 적절한 데이터 충돌 확률을 보이면서도 질의 응답시간의 성능을 개선할 수 있는 새로운 데이터 할당 기법을 제안할 것이다. 본 논문의 구성은 다음과 같다. 2장에서 기존에 개발된 기법에 대하여 기술하고, 3장에서는 본 논문이 제안하는 새로운 데이터 할당 기법에 대하여 기술한다. 4장에서는 제안된 기법의 성능을 평가, 분석하고, 마지막으로 5장에서 결론을 맺는다.

II. 관련 연구

이 장에서는 기존에 제안된 기법들에 대하여 기술한

다. 데이터 할당 기법들은 질의가 요청하는 데이터의 수에 따라 단일 데이터 질의를 위한 기법들과 다중 데이터 질의를 위한 기법들로 분류할 수 있다. 단일 데이터 질의를 위한 기법에는 평형(FLAT)^[1], DP^[2], 그리고 NODA^[5] 등과 같은 기법들이 있는데, 이들은 모두 데이터의 액세스 확률에 기초하여 데이터를 분할하여 방송채널에 할당했다. 가장 단순한 평형(FLAT) 기법은 각 방송채널에 동일 개수의 데이터를 배정하고, 각 채널은 이들을 라운드-로빈 형태로 평형 방송한다. 평형 기법은 구현이 용이한 장점이 있지만, 모든 데이터들을 동일 빈도로 방송하기 때문에 평균 응답시간의 성능이 낮다는 단점이 있다.

DP 기법은 각 방송채널 별로 평형 방송을 하는 시스템 모델을 사용하는데, 이 시스템 모델 하에서는 최적 알고리즘이다. 그러나 이 기법은 시스템 모델의 제약으로 인하여 같은 방송채널에 할당된 데이터들을 동일 빈도로 방송하기 때문에 이론적 최적 성능을 실현하진 못했다. NODA 기법은 각 방송채널에서 방송할 데이터들을 결정하기 위해, 평균 응답시간의 이론적 하한 값에 기초하여 전체 데이터들을 방송채널 수 만큼의 그룹으로 분할한다. 그리고 각 그룹의 데이터들을 다시 확장된 동적 프로그래밍 알고리즘을^[4] 사용하여 서브채널 수 만큼의 그룹으로 분할하여 서브채널에 할당하고, 각 서브채널별로 평형 방송을 실시한다. NODA는 이와 같이 각 방송채널이 동일 대역폭을 갖는 다수 개의 서브채널로 구성된다라는 개념을 고안하여 단일 데이터 질의의 응답시간 측면에서 이론적 최적에 거의 필적하는 성능을 실현했다.

다중 데이터 질의 환경에서는 QEM(Query Expansion Method)^[8], GA(Genetic Algorithm)^[9], PBA^[6], MA2D^[7], 그리고 DAMA(Data Allocation using Moving Average)^[10] 등과 같은 기법들이 제안되었다. QEM 기법은 단일채널 환경에서 방송 스케줄을 구성하는 문제를 연구한 반면, 다른 기법들은 다중 방송채널에 데이터를 할당하는 문제를 연구했다. QEM 기법은 질의에 포함된 데이터들의 응집성(coherence)을 표현하는 QD(Query Distance)를 정의하였는데, 어떤 질의의 QD 값이 작을수록 그 질의의 평균 응답시간은 작아지게 된다. 따라서 QEM 기법은 특정 질의에 포함된 데이터들을 방송 스케줄에 포함시킬 때 QD 값이 최소가 되는 방향으로 배치하는데, 질의 빈도(query frequency)에 우선권을 주어 질의 빈도가 높은 질의에 포함된 데이터들을 먼저 방송 스케줄에 배치한다.

GA 기법은 유전자 알고리즘을 이용하여 다중채널에

데이터들을 할당하는 문제를 연구했다. 이 기법은 반복적인 적합성 평가를 통하여 최적의 방송 스케줄을 찾아나가는데, 세대가 진행될수록 더 바람직한 해가 구해지긴 하지만, 이것이 최적이라는 보장도 없고 우수한 해를 구하기 위해서는 많은 세대를 거쳐야 하기 때문에 많은 실행 시간을 필요로 한다는 단점이 있다.

일반적으로 클라이언트는 한 순간에는 한 방송채널만을 수신하는데, 한 질의에 포함된 데이터가 동시에 방송되면 이 중 하나를 수신하고 다른 데이터는 다음 방송에서 수신해야 한다. 이와 같이 한 질의에 속한 데이터가 동시에 방송되는 것을 데이터 충돌이 발생한다고 표현하는데, 어떤 질의를 처리할 때 데이터 충돌이 발생하면 응답시간이 매우 길어지게 된다. PBA 기법은 이 점을 고려하여 데이터 충돌을 최소화하기 위한 방법을 제시했다. PBA는 더 높은 빈도로 요청되는 질의에 속한 데이터들을 데이터 충돌이 발생하지 않도록 우선 할당하여 해당 질의에서의 데이터 충돌 가능성을 최소화시킴으로써 질의 평균 응답시간 성능을 향상시켰다. 그러나 PBA 기법은 한 방송주기에서 모든 데이터들을 한번만 할당하는 정책을 사용하였기 때문에 응답시간을 개선하는 데는 한계가 있었다. MA2D 기법은 클라이언트에서 상대적으로 자주 요청하는 데이터들을 한 방송주기에서 여러 번 할당하는 정책을 사용하여 이 단점을 개선했다. MA2D 기법은 질의 프로파일과 요청 확률을 이용하여 한 방송주기에서 각 데이터들의 할당 횟수를 결정한다. 그런데 이 기법은 데이터 요청 횟수에 너무 큰 가중치를 부여하는 경향이 있기 때문에 요청 확률이 높은 질의의 응답시간은 대폭 개선할 수 있지만, 요청 확률이 낮은 질의의 응답시간이 너무 길어진다는 단점이 있다.

DAMA 기법은 다른 기법들과는^[6-9] 다르게 질의 프로파일과 요청 확률을 정형화하기 어려운 환경에서 각 데이터의 액세스 확률만을 이용하여 다중 데이터 질의를 처리하기 위한 연구를 수행했다. 다중 데이터 질의는 가장 늦게 수신된 데이터에 의해 응답시간이 결정된다. 따라서 질의 내의 어느 한 데이터가 방송 빈도가 낮아 액세스하는데 시간이 많이 걸린다면, 다른 데이터를 아무리 빨리 수신하더라도 응답시간 측면에서는 도움이 되지 않는다. 이와 같은 점을 고려한다면, 액세스 확률이 높을수록 더 빈번하게 방송되어야 하지만 질의가 액세스하는 데이터의 수가 많아질수록 액세스 확률이 방송 빈도에 미치는 영향력이 작아져야 한다. DAMA 기법은 이 점에

착안하여 각 데이터의 액세스 확률의 이동 평균(Moving Average) 값을 이용하여 방송채널에 데이터들을 할당한다. 비록 DAMA 기법이 질의 크기와 데이터의 액세스 확률에 따라 각 데이터의 방송빈도를 적절하게 제어하지는 않지만, 질의 프로파일과 요청 확률을 이용한다면 그 성능을 더욱 개선시킬 수 있을 것이다.

III. 새로운 상호 관련 데이터 할당 기법

이 장에서는 본 논문이 제안한 새로운 데이터 할당 기법에 대하여 기술하는데, 제안된 기법은 상호 관련있는 데이터들의 특성을 이용하여 방송채널에 할당하기 때문에 IDAS(Interdependent Data Allocation Scheme)라 명명한다.

1. 기본 시스템 모델 및 용어

본 논문은 다음과 같은 시스템 모델과 용어를 사용한다.

1) 전체 데이터들은 $d_k(1 \leq k \leq N)$ 로 표현되는 동일 크기의 N 개의 데이터들로 구성되는데, 이들 중 질의에 포함된 데이터들만이 방송된다.

2) 전체 질의들은 $q_j(1 \leq j \leq m)$ 로 표현되는 m 개의 질의로 구성되는데, 각 질의의 요청 확률은 $p(q_j)$ 로 표현하며 내림차순으로 정렬되어 있다고 가정한다.

3) 방송 서버에는 동일 대역폭을 갖는 K 개의 방송채널이 존재하는데 이를 $CH_i(1 \leq i \leq K)$ 로 표현하며, 각 방송채널은 자신에게 배치된 데이터들을 순서대로 라운드-로빈 방식으로 주기적으로 방송한다.

4) 클라이언트는 어느 한 순간에는 한 방송채널만을 수신할 수 있다. 만일 어떤 질의에 포함된 데이터들이 동시에 방송되면 이 중 하나의 데이터만을 수신하며, 이 경우를 데이터 충돌이 발생했다고 표현한다.

5) 한 방송채널에서 하나의 데이터를 방송하는데 걸리는 논리적인 시간을 1 tick으로 정의하고, 방송주기 및 응답시간 성능의 단위로서 tick을 사용한다.

6) 각 방송채널에는 동일한 수의 데이터가 배치되는데, 만일 방송채널에 T 개의 데이터가 할당되었다면 방송주기는 T 이다. 방송채널에서 각 데이터가 방송되는 구간을 슬롯(slot)으로 표현하는데, 각 슬롯은 슬롯번호로서 구분한다. $Slot_{i,j}(1 \leq i \leq K, 1 \leq j \leq T)$ 는 CH_i 의 j 번째 슬롯을 의미한다.

2. 기본 개념 및 데이터 할당 알고리즘

PBA 기법은 요청 확률이 높은 질의가 액세스하는 데이터들에 우선권을 주어 충돌이 발생하지 않도록 방송채널에 할당하기 때문에 매우 낮은 데이터 충돌율을 보인다. 그러나 PBA 기법은 한 방송주기에서 모든 데이터들을 한번만 할당하는 정책을 사용하기 때문에 데이터 충돌이 발생하지 않는 질의들은 그들의 요청 빈도에 관계없이 모두 동일한 평균 응답시간을 보인다. 전체 평균 응답시간의 성능은 식 (1)과 같이 표현되는데 이를 향상시키기 위해서는 요청 확률이 높은 질의일수록 더 빠른 응답시간의 성능을 발휘하도록 데이터를 할당해야 한다. 이 수식에서 $AvgResTime(q_j)$ 는 질의 q_j 의 평균 응답시간을 의미한다.

$$\sum_{j=1}^m (p(q_j) \times AvgResTime(q_j)) \quad (1)$$

만일 요청 확률이 높은 질의에 포함된 데이터들을 한 방송주기에서 여러 번 할당한다면 이 질의의 평균 응답시간 성능은 대폭 개선될 것이다. 그러나 필요 이상으로 동일 데이터를 너무 많이 할당하면 요청 확률이 낮은 질의의 평균 응답시간이 너무 길어져 오히려 전체 평균 응답시간 성능은 저하될 것이다. 본 논문은 이와 같은 점을 고려해서 질의 프로파일과 요청 확률을 취합하여 각 데이터들의 방송 빈도를 적절히 설정하여 응답시간 성능을 개선할 수 있는 새로운 데이터 할당 기법을 제안한다. 단일 데이터 질의에서 각 데이터들의 상대적 방송빈도를 액세스 확률의 제곱근 값에 비례하게 설정할 때 가장 최적의 응답시간 성능을 발휘한다^[11]. 본 논문에서 이를 고려하여 각 데이터의 상대적인 방송 빈도를 액세스 확률의 제곱근 값에 비례하게 할 때 최적의 응답시간 성능을 보일 것이라는 휴리스틱을 적용하여 데이터들을 방송채널에 할당한다. 이를 위해 우선 각 질의 프로파일과 요청 확률로부터 각 데이터의 상대적인 액세스 확률 $RelAccP(d_i)$ 를 식 (2)와 같이 계산한다.

$$RelAccP(d_i) = \sum_{j=1}^m p(q_j) \quad (2)$$

그리고 각 데이터의 방송빈도를 상대적인 액세스 확률의 제곱근 값에 비례하게 설정하기 위하여, 한 방송주기에서 각 데이터의 할당 횟수를 식 (3)과 같이 결정한다. 이 수식에서 $MinRAP$ 은 방송되는 데이터들 중에서 액세스 확률이 가장 작은 데이터의 $RelAccP(d_i)$ 값을 의미한다

다. 이 수식에 의하면 한 방송주기에서 가장 액세스 확률이 작은 데이터를 한번 할당하고, 이를 기준으로 해서 다른 데이터들은 상대적인 액세스 확률의 제공급 값에 비례하게 할당 횟수를 결정한다. 한 방송주기에서 방송되는 데이터의 총 개수는 $\sum_{i=1}^N BCastFreq(d_i)$ 로 계산되고 방송채널의 수가 K 개이므로 방송주기 T 는 식 (4)와 같이 결정된다. 어떤 데이터를 한 방송주기에서 여러 번 방송할 때 효율성을 높이기 위해서는 그 데이터의 방송 간격을 일정하게 하는 것이 최적 성능을 발휘한다. 이런 특성을 반영하기 위해 각 데이터의 할당 간격은 식 (5)와 같이 결정한다.

$$BCastFreq(d_i) = \lfloor \sqrt{RelAccP(d_i)} / \sqrt{MinRAP} + 0.5 \rfloor \quad (3)$$

$$T = \lceil (\sum_{i=1}^N BCastFreq(d_i)) / K \rceil \quad (4)$$

$$BCastInt(d_i) = \lceil T / BCastFreq(d_i) \rceil \quad (5)$$

질의 프로파일과 요청 확률에 기초하여 질의가 액세스하는 데이터들을 방송채널에 할당하는 전체 알고리즘이 그림 1에 제시되어 있다. 이 알고리즘은 우선 식(2)~(5)를 이용하여 한 방송주기에서 각 데이터의 할당 횟수 및 할당 간격을 결정한다. 제안된 IDAS 알고리즘은 요청 확률이 높은 질의일수록 데이터 충돌 가능성을 줄이기 위하여, 요청 확률이 높은 질의의 순서대로 우선권을 주어 각 질의의 데이터들을 방송채널에 할당한다. IDAS는 *ScheduledDataSet*을 유지 관리하는데, 어떤 데이터가 방송채널에 할당될 때마다 *ScheduledDataSet*에 이를 추가한다. K 개의 방송채널이 존재하기 때문에 어떤 슬롯번호에는 최대 K 개의 데이터들이 할당되는데, 더 이상 데이터 할당이 불가능한 슬롯 번호를 저장하기 위해 *FullSlotSet*을 유지 관리한다.

알고리즘 이해를 위해 질의 q_1, q_2 는 각각 $\{a, b, c\}, \{x, y\}$ 를 액세스 하고 $\{a, b, c\}$ 의 방송 횟수는 3, $\{x, y\}$ 의 방송 횟수는 2로 가정하자. 그리고 방송주기는 12라고 가정하고 그림 2를 살펴보자. 식 (5)에 의해 $\{a, b, c\}$ 의 방송 간격은 4, $\{x, y\}$ 의 방송 간격은 6이 될 것이다. 알고리즘 ①번 부위에 의해 $CurChNo$ 와 $CurSlotNo$ 가 모두 1로 결정되어 a 는 $Slot_{1,1}$ 에 할당된다. 그리고 알고리즘 ②번 부위에 의해 $CurChNo$ 를 다음 방송채널로 이동하고 ③번 부위에 의해 $CurSlotNo$ 를 우측으로 방송간격 4만큼 이동한다. 따라서 a 의 다음 할당 위치는 $Slot_{2,5}$ 로 결정되고 그 다음 할당 위치는 $Slot_{3,9}$ 로 결정된다. a 가 전부 할당되고 나면 b 를 할당하기 시작하는데 b 의 처음 할당 위치가

①번 부위에 의해 $Slot_{1,2}$ 로 결정되고 다음 할당 위치는 $Slot_{2,6}$ 과 $Slot_{3,10}$ 으로 결정된다. 마찬가지로 방법으로 c 의 할당 위치는 그림 2와 같이 결정된다. q_1 의 데이터가 모두 할당되고 나면, q_2 에 포함된 데이터들이 할당되는데 우선 x 의 처음 할당 위치가 ①번 부위에 의해 $Slot_{1,4}$ 로 결정되고, 다음 할당 위치는 $Slot_{2,10}$ 으로 결정된다. y 의 할당 위치는 그림 2와 같이 결정된다.

```

//N: number of data items, m: number of queries,
//K: number of broadcast channels, T: broadcast period
RelAccP(d_i) = \sum_{j=1}^m \sum_{d_i \in q_j} (p(q_j))
SqrtRAP(d_i) = \sqrt{RelAccP(d_i)}
MinSqrtRAP = \min(SqrtRAP(d_i)) (1 \le i \le N)
BCastFreq(d_i) = \lfloor SqrtRAP(d_i) / MinSqrtRAP + 0.5 \rfloor
BCastDataCount = \sum_{i=1}^N (BCastFreq(d_i))
T = \lceil BCastDataCount / K \rceil
BCastInt(d_i) = \lceil T / BCastFreq(d_i) \rceil
//Allocate unscheduled data items of each query
set ScheduledDataSet and FullSlotSet empty.
for each query q_i from i=1 to i=m
  for each data item d_k in query q_i
    if (d_k is not in ScheduledDataSet)
      ① find an empty slot Slot_{CurChNo, CurSlotNo} such that
         CurChNo is minimized and then CurSlotNo is
         minimized.
         add d_k in the ScheduledDataSet.
         loop BCastFreq(d_k) times
         allocate d_k into Slot_{CurChNo, CurSlotNo}.
         if (K data items are allocated in CurSlotNo)
           add CurSlotNo in the FullSlotSet.
         end if
      ② CurChNo = (CurChNo mod K) + 1
      ③ shift CurSlotNo to the right by BCastInt(d_k)
         in round-robin manner.
         if (CurSlotNo is not an element of FullSlotSet)
           if (Slot_{CurChNo, CurSlotNo} is empty)
             continue
           end if
         ④ find x such that Slot_{x, CurSlotNo} is empty, and
            set x into CurChNo.
         else
           ⑤ find the nearest empty slot to the CurSlotNo,
              and set its slot number and channel number
              into CurSlotNo and CurChNo, respectively.
         end if
         end loop
       end if
     end for
   end for

```

그림 1. 데이터 할당 알고리즘
 Fig. 1. Data allocation algorithm

slot no.	1	2	3	4	5	6	7	8	9	10	11	12
CH ₁	a	b	c	x	y							
CH ₂					a	b	c			x	y	
CH ₃									a	b	c	

그림 2. 데이터 할당의 예
Fig. 2. Example of data allocation

그림 2에서 보는 것과 같이 우선 순위가 높은 질의에 포함된 데이터들의 방송 횟수가 동일하다면 이들은 대부분 충돌이 발생하지 않는다. 뿐만 아니라 동일 질의에 포함된 데이터들이 서로 인접해서 할당되기 때문에 우수한 평균 응답시간의 성능을 보일 것이다. 그리고 요청 확률이 높은 질의의 데이터들은 액세스 확률이 높아 다수 번 방송채널에 할당되기 때문에 설사 충돌이 발생하더라도 한 방송주기를 기다리는 것이 아니라 자신의 방송간격만큼만 기다리면 되므로 데이터 충돌로 인한 부담이 적은 편이다. 그러나 요청 확률이 낮은 질의의 경우에는 자신의 데이터가 할당되어야 위치에 이미 다른 데이터들이 할당되어 있을 가능성이 커지기 때문에 데이터 충돌 가능성이 높아지게 된다.

만일 $Slot_{CurChNo, CurSlotNo}$ 에 이미 다른 데이터가 할당되어 있다면 $CurSlotNo$ 에서 빈 슬롯을 가진 방송채널을 찾아 $CurChNo$ 를 다시 설정하고 해당 슬롯에 데이터를 할당한다(그림 1의 ④번 부위 참조). 만일 $CurSlotNo$ 에 이미 K 개의 데이터들이 할당되어 있다면 $CurSlotNo$ 에서 가장 근접한 빈 슬롯을 찾아 $CurChNo$ 와 $CurSlotNo$ 를 다시 설정하여 다음번 할당 위치를 결정한다(그림 1의 ⑤번 부위 참조). 한 방송주기에서 방송되는 총 데이터의 개수가 K 의 정수배가 아니면, 그림 1의 알고리즘을 실행한 후에도 빈 슬롯이 존재한다. 이 경우에는 임의의 데이터를 빈 슬롯에 할당한다.

IV. 성능 평가 모델 및 성능 평가

이 장에서는 본 논문이 제안한 IDAS 기법의 성능을 평가하고 그 결과를 기술한다. 비교 대상으로는 FLAT^[1], PBA^[6], 그리고 DAMA^[10] 기법을 선정하였다.

1. 성능평가 모델

본 논문의 성능평가 모델은 크게 K 개의 방송채널 서버와 모바일 클라이언트로 구성된다. 클라이언트는 질의 확률 모델에 따라 질의를 제기하고 방송채널을 감시하다

가 질의에 포함된 데이터가 방송되면 이를 수신한다. 클라이언트는 현재 질의가 필요로 하는 모든 데이터를 수신한 후에 다음 질의를 제기하며, 클라이언트는 한 순간에는 한 방송채널만 수신할 수 있다. 본 성능평가 모델은 MCC에서 개발한 CSIM^[12]을 이용하여 구현하였으며, 본 논문에서 사용된 입력변수들은 표 1과 같다. m 은 클라이언트가 제기할 수 있는 질의 종류의 개수이며, Q 는 질의 크기를 의미하는데 실제 질의가 액세스하는 데이터의 수는 Q 에서 $\pm 20\%$ 편차를 가지게 무작위로 결정된다.

표 1. 성능 평가를 위한 입력 변수

Table 1. Input parameters for performance evaluation

입력변수	의미	기본값	설정
N	데이터 항목의 수	5000	5000
K	방송채널의 수	10	1 ~ 20
m	질의의 개수	200	200
Q	질의의 평균 크기	10	1 ~ 20
θ	zipf 분포의 θ 값	0.95	0.0 ~ 1.4

클라이언트가 제기하는 질의들은 불균등한 액세스 형태를 모델링하기 위해 많이 사용되는 zipf 분포 모델^[11] 따르다고 가정한다. zipf 분포 모델에서 각 질의 요청 확률은 식 (6)으로 표현되는데, 여기서 θ 값은 각 질의의 요청 확률이 어느 정도 편향되어 있는지를 나타내는 인수로써, θ 값이 0일 경우에는 각 질의의 요청 확률은 모두 동일하며 θ 값이 커질수록 각 질의들의 요청 확률간의 차이가 커지게 된다. 방송 서버의 데이터베이스는 동일한 크기를 가진 N 개의 데이터들로 구성되는데, 각 질의가 액세스하는 데이터들은 전체 N 개의 데이터들 중에서 임의로 결정된다. N 개의 데이터들 중에서 질의에 포함된 데이터들만이 각 기법에 맞추어 방송채널에 할당되어 주기적으로 방송된다.

$$p(q_i) = \frac{(1/i)^\theta}{\sum_{j=1}^m (1/j)^\theta} \quad (1 \leq i \leq m) \quad (6)$$

데이터 할당 기법의 주요 평가기준은 질의의 평균 응답시간이므로 본 논문은 이를 중심으로 성능을 평가한다. 성능평가 실험에서 실험 시작시의 초기 편향(initial bias)을 제거하기 위하여 초기 십만 개 질의들의 실험 결과는 무시하였다. 성능 결과에서 제시된 값은 5 개의 서로 다른 임의의 수를 사용하여 실시된 실험 결과의 평균값으로, 각 실험은 천만 개의 질의들이 처리될 때까지 실시하였다.

2. 질의 크기에 따른 평균 응답시간의 성능

본 실험에서는 한 질의가 액세스하는 데이터의 수를 의미하는 질의 크기를 1에서 20까지 변화시키면서 각 기법의 성능추이를 살펴보았다. 방송채널의 수는 10으로 설정했으며, θ 값은 데이터의 편향적인 액세스를 모델링하기 위한 대표 값으로 많이 사용되는 0.95로 설정하였다^[1]. 질의의 평균 응답시간 성능이 그림 3에 제시되어 있는데, 대체로 IDAS, DAMA, PBA, 그리고 FLAT 기법 순으로 우수한 성능을 보인다. 그래프 상으로는 구분이 어렵지만, 질의 크기가 1~4 구간에서도 IDAS는 다른 기법보다 우수한 성능을 보인다.

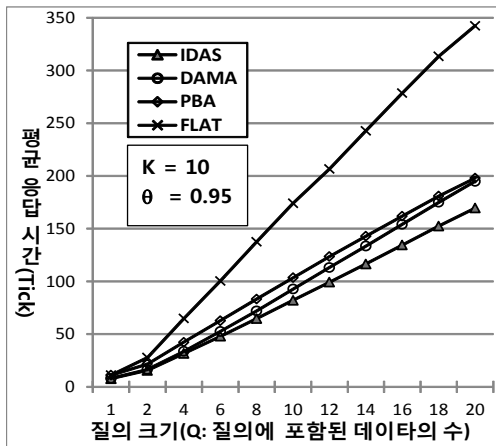


그림 3. 질의 크기에 따른 평균 응답시간
 Fig. 3. Query size vs. average response time

IDAS가 다른 기법에 비해 꾸준히 우수한 성능을 보이는데 이는 다음의 두 가지 원인에서 기인된다. 첫째, IDAS 기법은 상대적 액세스 확률의 제곱근 값에 비례하게 각 데이터의 방송 횟수를 결정하는 정책을 사용하여, 요청 확률이 높은 질의일수록 빠르게 처리되는 특성을 가지고 있다. 둘째, IDAS는 요청 확률이 높은 질의 순서대로 그 질의에 포함된 데이터들을 방송채널에 할당하기 때문에 요청 확률이 높은 질의의 데이터들은 서로 인접하게 할당될 가능성이 높게 된다. 그리고 이들은 비교적 데이터 충돌률도 작은 편이다.

Q가 증가함에 따라 모든 기법들의 성능은 점점 저하되는데, 이는 질의를 처리하기 위하여 방송채널로부터 더 많은 수의 데이터를 수신해야 할 뿐만 아니라 데이터 충돌 확률도 점점 더 커지기 때문이다. FLAT 기법은 질의 프로파일에 관계없이 방송되는 데이터들을 일정한 수

로 분할하여 각 방송채널에 할당하므로 모든 데이터들은 동일 빈도로 방송한다. 그리고 질의에 포함된 데이터들의 인접 할당 및 데이터 충돌 회피를 위한 어떤 수단도 가지고 있지 않기 다른 기법에 비해 매우 낮은 성능을 보인다. PBA도 FLAT 기법과 같이 모든 데이터들을 동일 빈도로 방송한다. 그렇지만 PBA 기법은 요청 확률이 높은 질의에 우선권을 주어 이들의 충돌 가능성을 최소화시켰고 질의에 포함된 데이터들을 인접하게 할당하기 때문에 FLAT 기법보다는 훨씬 우수한 응답시간의 성능을 발휘한다.

DAMA 기법에서도 모든 데이터들은 방송채널에 한 번만 할당된다. 그러나 DAMA는 다른 기법과는 다르게 각 방송채널의 방송주기가 다르다. DAMA 기법은 액세스 확률이 높은 데이터들이 할당되는 방송채널에는 적은 수의 데이터들을 할당하는 정책을 사용하여, 액세스 확률에 따라 방송빈도를 다르게 할 수 있다. 그러나 동일 채널에 할당된 데이터들은 동일 빈도로 방송해야 한다는 한계가 있기 때문에 IDAS처럼 세밀하게 방송 빈도를 결정하지는 못한다. DAMA 기법은 데이터 충돌 회피를 위한 수단을 가지고 있지 않지만 동일 질의에 포함된 데이터들은 액세스 확률이 유사할 가능성이 크기 때문에 동일 방송채널에 할당될 가능성이 높게 된다. 따라서 DAMA 기법은 FLAT 기법보다는 더 낮은 데이터 충돌률을 보이게 된다.

질의 데이터의 인접성 측면에서는 PBA, IDAS, DAMA, FLAT 순으로 우수한 편이다. 비록 IDAS가 PBA 기법보다 데이터 인접성은 떨어지지만, IDAS는 PBA 기법과는 다르게 각 데이터의 방송 빈도를 다르게 설정할 수 있다. 따라서 IDAS는 전체 평균 응답시간의 성능에 큰 영향을 미치는 요청 확률이 높은 질의들을 매우 빠르게 처리할 수 있기 때문에 PBA 기법보다 우수한 성능을 보인다.

3. 채널 수의 변화에 따른 평균 응답시간의 성능

이 절에서는 방송채널의 수를 1~20으로 변화시키면서 각 기법의 성능 추이를 살펴보았다. 질의 크기를 10으로 설정하고 θ 값은 0.95로 설정하고 실시한 평균 응답시간 성능이 그림 4에 제시되어 있다. 이 시험에서도 본 논문이 제안한 IDAS는 다른 기법에 비해 전 채널 구간에서 훨씬 우수한 성능을 보인다. 그림 4를 살펴보면 모든 기법에서 K가 증가함에 따라 평균 응답시간이 감소하는 것

을 확인할 수 있다. 이는 방송채널 수가 증가함에 따라, 각 채널에 할당되는 데이터 개수가 작아져 데이터들을 더 빈번하게 방송할 수 있기 때문이다.

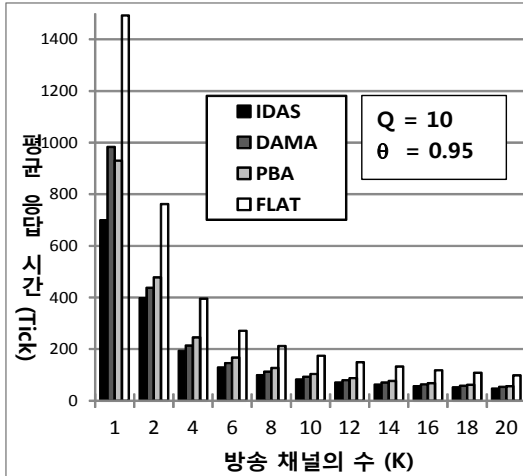


그림 4. 방송채널의 수에 따른 평균 응답시간
Fig. 4. No. of channel vs. average response time

방송채널의 수가 1인 경우에는 PBA, DAMA, 그리고 FLAT 기법은 모든 방송 데이터들을 동일 빈도로 방송하며 데이터 충돌도 전혀 발생하지 않는다. 따라서 이들은 K가 1이면 서로 동일한 성능을 보일 것 같지만 실제로는 그렇지 않다. 이는 전적으로 질의가 액세스하는 데이터들이 방송채널에 인접해서 배치되어 있나 여부에 기인한다. 이에 대한 이해를 돕기 위해 어떤 방송채널에 400개의 데이터가 할당되어 있다고 하고, 질의 q_1 이 $\{d_{101}, d_{201}, d_{301}, d_{401}\}$ 를 액세스한다고 가정하자. 만일 이 데이터들이 슬롯번호 {1, 2, 3, 4}와 같이 서로 인접해서 할당되어 있다면, q_1 의 평균 응답시간은 대략 방송주기의 반인 200 tick이 될 것이다. 그러나 이 데이터들이 슬롯번호 {1, 101, 201, 301}에 할당되어 있다면 q_1 은 350 tick 정도의 평균 응답시간을 보일 것이다.

PBA 기법은 질의 순으로 질의에 포함된 데이터들을 방송채널에 할당하기 때문에 동일 질의의 데이터들은 서로 인접해서 할당되는 특성이 있다. DAMA 기법은 액세스 확률이 높은 데이터 순으로 배치하는데, 같은 질의에 포함된 데이터들은 액세스 확률이 유사할 가능성이 크기 때문에 질의의 데이터들이 어느 정도 인접해서 할당된다. 그러나 FLAT 기법은 데이터들을 자신의 인덱스 순으로 방송채널에 할당하기 때문에 질의 데이터들의 인접성이

매우 낮은 편이다. IDAS는 PBA 기법처럼 요청 확률이 높은 질의 순으로 그 질의에 포함된 데이터들을 방송채널에 할당하기 때문에 질의 데이터의 인접성이 우수한 편이다. 그리고 방송채널이 1 개만 존재할 때도 데이터들의 액세스 확률을 고려하여 방송 빈도를 다르게 설정할 수 있다. 이런 이유들로 인해서 K가 1인 경우에는 IDAS, PBA, DAMA, 그리고 FLAT 기법 순으로 우수한 성능을 보이는 것이다.

그러나 K가 2 이상으로 증가하면 DAMA는 PBA 기법보다 더 우수한 성능을 보이기 시작한다. 다수의 방송채널이 존재하면 데이터 충돌이 발생한다. PBA는 DAMA 기법에 비해 데이터 충돌률과 질의 데이터들의 인접성 측면에서 더 우수한 특성을 가지고 있다. 반면 DAMA 기법은 데이터 액세스 확률에 따라 방송빈도를 다르게 할 수 있다는 특성이 있다. 다수의 방송채널이 존재할 경우에는 DAMA의 장점이 PBA 장점보다는 더 크게 성능에 영향을 미치기 때문에 DAMA가 PBA 기법보다 더 우수한 성능을 보이는 것이다. 본 연구에서는 질의 크기를 5로 설정하고 방송채널의 수를 변경하면서 성능 추이를 살펴보는 실험을 실시했는데, 이 실험에서도 각 기법의 상대적인 성능이 그림 4와 거의 유사한 상대적인 평균 응답시간의 성능을 보인다.

4. θ 값의 변화에 따른 평균 응답시간의 성능

본 실험에서는 θ 값의 변화에 따른 평균 응답시간의 성능추이를 살펴보았다. K를 10으로, Q를 10으로 설정하고 θ 값을 변화시키면서 살펴본 성능 결과가 그림 5에 제시되어 있다. θ 값이 0이 되면, 모든 질의의 요청 확률이 동일하게 되는데, 이것이 각 데이터의 액세스 확률이 동일하다는 것을 의미하지는 않는다. 왜냐하면 각 질의 프로파일을 생성할 때 질의가 액세스하는 데이터가 무작위로 선정되기는 하지만, 한 데이터가 여러 질의에 포함될 수도 있기 때문이다. 이 실험에서는 θ 값이 증가하게 되면 FLAT을 제외한 다른 기법들은 평균 응답시간의 성능이 급격하게 향상된다. PBA 기법의 경우에는 요청 확률이 높은 질의의 데이터 인접성과 데이터 충돌률이 매우 우수하는데, θ 가 증가하면 전체적인 데이터 인접성이 향상되고 데이터 충돌률도 감소하기 때문에 성능이 향상되는 것이다. 질의 요청 확률간의 차이가 커지게 되면 동일 질의에 속한 데이터의 액세스 확률은 그 질의의 요청 확률에 비례하여 증가하거나 감소하게 된다. IDAS와

DAMA 기법의 경우에는 각 데이터의 액세스 확률에 따라 각 데이터의 방송 빈도를 적절하게 설정하기 위한 매커니즘을 가지고 있기 때문에 θ 가 증가하면 평균 응답시간의 성능이 향상되는 것이다. FLAT 기법은 데이터를 방송채널에 데이터의 인덱스 순으로 배치하기 때문에 θ 값이 증가하더라도 성능이 미미하게 향상될 뿐이다.

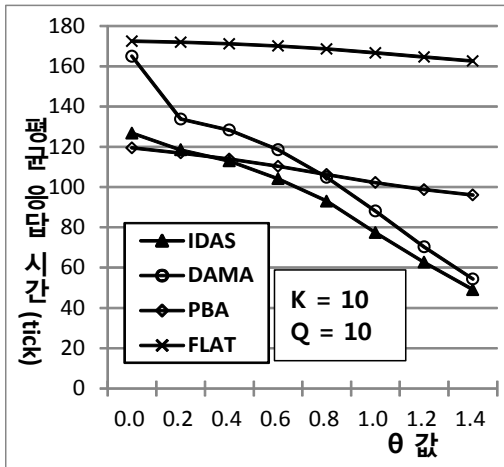


그림 5. θ 값의 변화에 따른 평균 응답시간
 Fig. 5. θ value vs. average response time

θ 값이 작으면 각 질의에 포함된 데이터들의 대부분은 동일한 액세스 확률을 보이며, 일부만이 더 높은 액세스 확률을 보인다. 예를 들어, 이 환경에서 방송되는 총 데이터의 수는 대략 1650 개 정도인데, 이들 중에서 한 질의에만 포함된 데이터는 1340 개, 두 질의에 포함된 데이터는 264 개, 세 개 이상의 질의에 속한 데이터는 46 개 정도가 된다. 만일 θ 값이 0일 경우에는 모든 질의의 요청 확률이 동일하므로 80% 이상의 데이터들은 동일한 액세스 확률을 보이게 된다. 이는 80% 이상의 데이터들이 한 방송주기에서 한번 방송됨을 의미한다. 이와 같은 경우에는 평균 응답시간은 이들 데이터들에 의해 결정될 가능성이 높기 때문에 IDAS의 데이터 액세스 확률에 기초하여 방송 빈도를 다르게 설정하는 정책은 오히려 방송주기만 길어지게 하고 데이터의 인접성만 훼손하는 결과를 가져오게 된다. DAMA 기법은 액세스 확률이 높은 데이터들이 배정되는 방송채널에는 적은 수의 데이터를 할당함에 의해 데이터들의 방송빈도를 다르게 설정한다. 뿐만 아니라 이 기법은 질의 순서대호가 아니라 데이터의 액세스 확률 순으로 방송채널에 데이터를 할당하기

때문에 IDAS나 PBA 기법보다 데이터 인접성이 나쁜 편이다. 특히 θ 가 작을 경우에는 데이터의 인접성이 급격하게 훼손된다. 이와 같은 이유들로 인해서 이 실험에서 θ 값이 작을 경우는 PBA가 IDAS와 DAMA 기법보다 더 우수한 성능을 보이는 것이다.

그러나 θ 값이 증가하게 되면 성능 역전 현상이 발생하는데, IDAS는 θ 가 0.4 이상이 되면 PBA 기법보다 우수한 성능을 보이기 시작하며, DAMA는 θ 가 0.8 이상이 되면 PBA 기법보다 우수한 성능을 보이기 시작한다. 이는 θ 값이 커지게 되면 동일 질의에 포함된 데이터들은 유사한 비율로 액세스 확률이 높아지게 되고 이를 반영하여 방송 빈도를 다르게 설정하는 정책이 효과를 발휘하기 때문이다. 그리고 DAMA 기법의 경우에는 θ 값이 증가하면 질의의 데이터 인접성도 개선되는 효과가 있기 때문이다.

V. 결 론

본 논문은 방송 방식의 데이터 전달 방법을 채택한 다중 방송채널 환경의 정보시스템에서 다중 데이터 질의를 효율적으로 처리하기 위한 연구를 수행하여 IDAS라 명명된 새로운 데이터 할당 기법을 제안했다. 본 논문이 제안한 기법은 질의 프로파일과 질의 요청 확률을 취합하여 각 데이터의 상대적 액세스 확률을 구한 후, 이 값의 제공급 값에 비례하게 각 데이터의 방송 빈도를 결정한다. 본 논문은 다중 데이터 질의를 위한 시뮬레이션 모델을 설계 구현하여 성능평가를 실시했는데, 제안된 기법은 대부분의 환경에서 다른 기법보다 우수한 평균 응답시간의 성능을 보였다. 이는 IDAS 기법은 요청 확률이 높은 질의일수록 빠르게 처리할 수 있을 뿐만 아니라, 질의 데이터의 인접성도 우수하기 때문이다.

본 논문은 질의 프로파일과 질의 요청 확률이 주어진 환경에서 미리 방송채널에 데이터를 할당한 후 이를 주기적으로 방송하는 시스템 모델을 사용했다. 본 논문의 미래 연구로서는 클라이언트에서 직접 서버로 질의 처리를 요청하고 서버는 이 정보를 바탕으로 동적으로 방송 스케줄을 구성하는 요청형 방송(on-demand broadcast) 시스템 모델을 사용하여 연구를 진행할 것이다.

References

- [1] S. Acharya, "Broadcast Disks: Dissemination-based Data Management for Asymmetric Communication Environments," Ph.D. thesis, Brown University, 1998.
- [2] W. G. Yee, S. Navathe, E. Omiecinski, and C. Jermaine, "Efficient Data Allocation over Multiple Channels at Broadcast Servers," IEEE Trans. on Computers, Vol. 51, No. 10, pp. 1231-1236, Oct. 2002.
- [3] S. Wang and H.L. Chen, "An $O(N \log K)$ Restricted Dynamic Programming Algorithm for Data Allocation over Multiple Channels," IEICE trans. on communications, Vol. E88-B, No. 9, pp. 3756-3764, Sep. 2005.
- [4] H.M. Kwon, "TLDP: A New Broadcast Scheduling Scheme for Multiple Broadcast-Channel Environments," The Journal of the Institute of Webcasting, Internet and Telecommunication, Vol. 11, No. 2, pp. 63-72, 2011.
- [5] H.M. Kwon, "A Near Optimal Data Allocation Scheme for Multiple Broadcast-Channel Environments," The Journal of the Institute of Webcasting, Internet and Telecommunication, Vol. 12, No. 1, pp. 17-27, 2012.
- [6] H.P. Hung, J.W. Huang, J.L. Huang, and M.S. Chen, "Scheduling dependent items in data broadcasting environments," ACM SAC 2006.
- [7] S.W. Park and S.W. Jung, "Interdependent Data Allocation a scheme over Multiple Wireless Broadcast Channels," Journal of KIISE : Database, Vol. 36, No. 1, pp. 30-43, 2009.
- [8] Y.D. Chung and M.H. Kim, "Effective Data Placement for Wireless Broadcast," Distributed and Parallel Databases, Vol 9, No. 2, 2001
- [9] J.L. Huang and M.S. Chen, "Dependent data broadcasting for unordered queries in a multiple channel mobile environment," IEEE Trans. on Knowledge and Data Engr., Vol. 16, No. 6, 2004.
- [10] H.M. Kwon, "A Broadcast Data Allocation Scheme for Multiple-Data Queries Using Moving Average of Data Access Probability," The Journal of the Institute of Webcasting, Internet and Telecommunication, Vol. 14, No. 5, pp. 35-43, 2014.
- [11] C. Hsu, G. Lee, A.L.P. Chen, "A near optimal algorithm for generating broadcast programs on multiple channels," Proc. ACM 10th Int'l Conf. CIKM, Atlanta, Georgia, pp. 303-309, 2001.
- [12] H. Schwetman, 'CSIM Users' Guide for Use with CSIM Revision 16', Microelectronics and Computer Technology Corporation, 1992.

저자 소개

권혁민(중신회원)



- 1984년 : 서울대학교 제어계측공학과 학사
- 1994년 : 한국과학기술원 정보및통신공학과 석사
- 1998년 : 한국과학기술원 정보및통신공학과 박사
- 1984년 ~ 1991년 : 대우전자 중앙연구소 컴퓨터개발부 선임연구원
- 1999년 ~ 현재 : 세명대학교 정보통신학부 교수
<주관심분야 : 트랜잭션 처리, 분산/병렬 데이터베이스, 모바일 컴퓨팅>

※ 이 논문은 2014학년도 세명대학교 교내학술연구비 지원에 의해 수행된 연구임