

<http://dx.doi.org/10.7236/JIIBC.2014.14.5.35>

JIIBC 2014-5-5

## 데이터 액세스 확률의 이동 평균을 이용한 다중 데이터 질의를 위한 방송 데이터 할당 기법

### A Broadcast Data Allocation Scheme for Multiple-Data Queries Using Moving Average of Data Access Probability

권혁민\*

Hyeokmin Kwon\*

**요약** 데이터 할당 기술은 데이터 방송 시스템의 성능을 향상시키기 위해서 필수적이다. 본 논문은 다중 데이터 질의가 제기되는 환경에서 방송채널에 데이터를 할당하는 주제를 연구하여 DAMA(Data Allocation using Moving Average)로 명명된 새로운 데이터 할당 기법을 제안한다. 제안된 기법은 각 데이터의 방송빈도를 자신의 액세스 확률의 이동 평균에 의해 결정하는 전략을 채택한다. DAMA는 질의의 크기에 따라 액세스 확률의 영향력을 적절하게 제어할 수 있기 때문에 질의 응답시간의 성능을 향상시킬 수 있다. 제안된 기법의 성능 평가를 위해 시뮬레이션이 수행되었다. 실험 결과에 따르면, 평균 응답시간의 성능에서 DAMA는 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 in the environment where multiple-data queries are submitted, and proposes a new data allocation scheme named DAMA. The proposed scheme employs the strategy that the broadcast frequency of each data is determined by the moving average of its access probability. DAMA could enhance the performance of query response time since it is capable of controlling the influence of access probability properly according to the query size. Simulation is performed to evaluate the performance of the proposed scheme. The simulation results show that the performance of DAMA is superior to other schemes in terms of the average response time.

**Key Words** : Data allocation, broadcast schedule, moving average, simulation

## 1. 서론

무선 이동통신 기술과 모바일 장치의 발달은 휴대용 무선 기기를 통하여 언제 어디서나 3G, 4G, 또는 IEEE 802.11과 같은 무선 네트워크에 접속하여 원하는 정보를 얻을 수 있게 되었다. 모바일 컴퓨팅(mobile computing)

환경에서는 서버에서 클라이언트로의 하향 대역폭이 클라이언트에서 서버로의 상향 대역폭에 비해서 매우 크다는 특징을 보인다<sup>[1]</sup>. 그리고 모바일 클라이언트는 배터리 용량의 제한으로 인하여 자신이 필요로 하는 데이터를 서버에 직접 요청하는 것이 부담이 될 수도 있다. 이와 같은 비대칭적 통신 대역폭과 제한된 배터리 용량으로

\*종신회원, 세명대학교 정보통신학부

접수일자 : 2014년 9월 12일, 수정완료 : 2014년 10월 2일

게재확정일자 : 2014년 10월 10일

Received: 12 September, 2014 / Revised: 2 October, 2014

Accepted: 10 October, 2014

\*Corresponding Author: hmkwon@semyung.ac.kr

School of Information and Communication System, Semyung University, Korea

인하여, 모바일 컴퓨팅 환경에서는 요청-응답(request-response) 방식보다는 방송 방식에 기초한 데이터 전달 방법이 훨씬 바람직하고 효과적인 데이터 전송 기술로 알려져 있다<sup>[1-2]</sup>. 이에 따라 모바일 클라이언트로 데이터를 효과적으로 방송하기 위한 많은 연구들이 수행되었다<sup>[1-12]</sup>. 방송 방식의 정보 시스템에서 방송 서버는 방송 채널을 통하여 순차적, 그리고 주기적으로 데이터들을 방송하고 모바일 클라이언트는 자신이 원하는 데이터가 방송 채널에 나타나면 이를 액세스한다.

그런데 방송 대역폭은 다수의 데이터들에 의해 공유되는 제한된 자원이기 때문에 방송 채널에 어떻게 데이터를 할당하고, 할당된 데이터들은 어떤 순서로 방송할 것인가를 결정하기 위한 데이터 할당 및 방송 스케줄링 기법은 데이터 방송 시스템에서 매우 중요한 기술이다. 질의의 응답 시간(response time)은 데이터 할당 및 방송 스케줄링 알고리즘의 성능을 평가하는 가장 중요한 요소로 사용되어 왔는데, 응답 시간이란 클라이언트가 특정 질의를 제기한 순간부터 자신이 원하는 모든 데이터들을 수신할 때까지 걸린 대기 시간을 의미한다. 그 동안 이 응답 시간의 성능을 향상시키기 위하여 단일 방송채널 환경에서 방송 스케줄링 기법에 관한 연구들이 수행되었는데, 최근에는 이를 다중 방송채널 환경으로 확장하는 연구들이 많이 수행되었다<sup>[3-7]</sup>. 다중 방송채널 환경에서는 각 채널의 방송 스케줄을 구성하는 문제와 더불어 전체 데이터들을 어떻게 분할하여 각 방송 채널에 할당한 것인가의 문제를 효율적으로 해결해야 한다. 데이터 할당 및 방송 스케줄링 기법들은 하나의 데이터를 액세스하는 단일 데이터 질의를 위한 기법들과 다수의 데이터를 액세스하는 다중 데이터 질의를 위한 기법들로 분류할 수 있다.

단일 데이터 질의 환경을 위한 많은 연구들이<sup>[3-7]</sup> 수행되었는데, 이들은 모두 데이터의 액세스 확률만이 주어진 환경에서 데이터를 분할하여 방송채널에 할당했다. DP(Dynamic Programming)<sup>[3]</sup>와 RDP(Restricted Dynamic Programming)<sup>[4-5]</sup> 기법은 각 방송채널이 평형(Flat) 방송된다는 가정 하에 응답시간이 최소화되도록 데이터를 분할하여 각 채널에 할당했다. 그런데 이 기법들은 같은 채널에 할당된 데이터들을 그들의 인기도에 상관없이 동일 빈도로 방송하기 때문에 응답시간의 성능을 향상시키는 데는 한계가 있었다. TLDP(Two-Level Dynamic Programming)<sup>[6]</sup>와 NODA(Near Optimal Data

Allocation)<sup>[7]</sup> 기법은 이 단점을 극복하기 위하여 각 방송 채널은 동일 대역폭을 갖는 다수의 서브채널로 구성되는데 개념을 도입하여 서브채널 별로 평형 방식으로 방송하는 시스템 모델을 정립하여 응답시간의 성능을 획기적으로 개선했다.

비록 TLDP와 NODA 기법이 이론적으로 최적 응답 시간에 거의 필적하는 성능을 보이기는 하지만, 이는 단일 데이터를 액세스하는 질의에 한한 것이다. 그러나 현실적으로는 다수 개의 데이터를 액세스하는 질의가 많이 제기되는 환경들이 존재한다. 예를 들어, 교통 정보시스템에서 교통 상황을 고려한 빠른 길을 찾기 위해서는 다수 개의 도로에 대한 교통 상황 정보를 검색하는 질의가 제기될 것이다. 이와 같은 다중 데이터 질의가 제기되는 환경을 위하여 다수의 데이터 할당 기법들이<sup>[8-11]</sup> 제안되었는데, 이들 기법들은 모두 질의에 포함된 데이터 집합을 의미하는 질의 패턴과 질의의 요청확률이 제시된 환경에서 이를 토대로 데이터를 방송채널에 할당하는 연구를 수행하였다. 그러나 질의 패턴과 질의 요청 확률을 취합하는 것은 매우 어려운 작업이며, 경우에 따라서는 이를 정형화하기 어려운 환경도 많이 있다. 예를 들어 증권 회사에서 전체 주식의 매매 빈도를 취합하면 전체 데이터의 액세스 확률은 쉽게 유추할 수 있지만, 이를 토대로 질의 패턴이나 질의 빈도를 유추하는 것은 쉽지가 않다.

본 논문은 이런 환경을 고려하여, 데이터들의 액세스 확률과 질의 크기만이 주어진 환경에서 다중 데이터 질의를 처리하기 위한 데이터 할당 기법에 관한 연구를 진행하여 질의의 응답시간 성능을 개선할 수 있는 새로운 기법을 제안한다. 본 논문의 구성은 다음과 같다. 2 장에서 기존에 개발된 기법에 대하여 기술하고, 3 장에서는 본 논문이 제안하는 새로운 데이터 할당 기법에 대하여 기술한다. 4장에서는 제안된 기법의 성능을 평가, 분석하고, 마지막으로 5 장에서 결론을 맺는다.

## II. 관련 연구

이 장에서는 기존에 제안된 기법들에 대하여 기술한다. 데이터 할당 기법들은 질의가 액세스하는 데이터의 수에 따라 단일 데이터 질의를 위한 기법들과 다중 데이터 질의를 위한 기법들로 분류할 수 있다. 단일 데이터 질의를 위한 기법에는 평형(FLAT)<sup>[1]</sup>, DP<sup>[3]</sup>, RDP<sup>[4-5]</sup>,

TLDP<sup>[6]</sup>, 그리고 NODA<sup>[7]</sup> 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 용이하다는 장점이 있지만, 모든 데이터들을 동일 빈도로 방송하기 때문에 평균 응답시간의 성능이 다른 기법들에 비하여 심각하게 저하되는 단점이 있다.

DP 기법은 각 채널의 데이터가 평형 방송된다는 가정 하에 평균 응답시간이 최소화되도록 전체 데이터들을 방송채널의 수만큼의 그룹으로 분할하여 이들을 각 방송채널에 할당한다. DP 기법은 각 채널의 데이터들이 평형 방송되는 시스템 모델에서 단일 데이터 질의의 응답시간 성능 측면에서는 최적(optimal) 알고리즘이다. 그러나 DP 기법은 가능한 모든 분할 방법을 고려하여 일일이 응답시간의 성능을 계산해야 하기 때문에 이의 실행을 위한 부담이 크다는 단점이 있다. Wang과 Chen의 연구에서는<sup>[4-5]</sup> DP의 수행시간 측면의 단점을 개선하기 위하여, 이론적으로 최적 응답시간의 성능을 추정하여 데이터를 분할하여 각 방송채널에 할당하는 RDP 기법을 제안했다.

DP와 RDP 기법은 각 방송채널이 평형 방송을 하는 시스템 모델에서는 최적 성능을 발휘하지만, 각 채널의 데이터들을 동일 빈도로 방송하기 때문에 응답시간의 성능을 개선하는 데는 한계가 있었다. 이 한계를 극복하기 위하여 TLDP와 NODA 기법은 각 방송채널이 동일 대역 폭을 갖는 다수 개의 서브채널로 구성된다는 개념을 도입하여 각 서브채널별로 평형 방송을 실시하는 시스템 모델을 정립했다. 이 기법들은 서브채널의 개념을 통하여, 동일 방송채널에 할당된 데이터들도 인기를 반영하여 다른 빈도로 방송할 수 있기 때문에 다른 기법들에<sup>[3-5]</sup> 비해 매우 우수한 응답시간의 성능을 발휘할 뿐만 아니라, 단일 데이터 질의의 응답시간 측면에서 이론적 최적 성능에 거의 필적하는 성능을 보인다.

다중 데이터 질의 환경을 위하여 QEM(Query Expansion Method)<sup>[8]</sup>, MQEM(Modified Query Expansion Method)<sup>[9]</sup>, GA(Genetic Algorithm)<sup>[10]</sup>, 그리고 PBA(Placement-Based Allocation)<sup>[11]</sup> 등과 같은 데이터 할당 기법들이 제안되었는데, 이들은 질의 패턴 및 질의 요청확률에 기초하여 방송채널에 데이터를 할당했다. QEM 기법은 단일 채널 환경에서 다중 데이터 질의를 처리하기 위한 방송 스케줄을 구성하는 문제를 연구했다. QEM 기법은 질의

에 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. QEM 기법은 질의에 포함된 데이터들을 방송 스케줄에 배치할 때 자신보다 먼저 방송 스케줄에 포함시킨 질의의 QD 값을 변경시키지 않았는데, MQEM 기법은 이 제약을 완화하는 정책을 사용하여 평균 응답시간 성능을 개선했다.

QEM과 MQEM 기법에서는 단일채널에 데이터를 할당하는 문제를 연구한 반면, GA와 PBA 기법의 연구에서는 다중 방송채널 환경도 고려했다. GA 기법은 유전자 알고리즘을 이용하여 다중채널에 데이터들을 할당했는데, 이 기법은 실행 부담이 매우 크다는 단점이 있다. PBA 기법은 질의에 포함된 데이터들이 동시에 같은 시간대에 방송됨으로써 질의 응답시간이 커지는 문제점을 해결하기 위한 연구를 하였다. PBA는 더 높은 빈도로 요청되는 질의에 속한 데이터들을 데이터 충돌이 발생하지 않도록 우선 할당하여 해당 질의에서의 데이터 충돌 가능성을 줄임으로써 질의의 평균 응답시간 성능을 향상시켰다.

### III. 새로운 방송 데이터 할당 기법

이 장에서는 본 논문이 제안하는 새로운 데이터 할당 기법에 대하여 기술한다. 기존에 제안된 다중 데이터 질의를 위한 데이터 할당 기법들은<sup>[8-11]</sup> 질의 패턴 및 질의 요청 확률이 제시된 환경에서 연구를 수행하였는데, 본 논문은 이런 정보를 취합하는 것이 어려운 환경을 고려하여 데이터들의 액세스 확률과 질의 크기만이 주어진 환경에서 다중 데이터 질의를 처리할 수 있는 데이터 할당 기법을 연구한다.

#### 1. 기본 시스템 모델

본 논문은 다음과 같은 시스템 모델과 용어를 사용한다.

1) 방송 데이터들은  $d_k(1 \leq k \leq N)$ 로 표현되는 동일 크기

의  $N$  개의 데이터로 구성되며,  $d_k$ 의 액세스 확률을  $p_k$ 로 표현한다. 그리고 데이터는 액세스 확률의 내림차순으로 정렬되어 있다고 가정한다.

- 2) 방송 서버에는 동일 대역폭을 갖는  $K$  개의 방송채널이 존재하는데, 이를  $CH_i(1 \leq i \leq K)$ 로 표현한다.
- 3) 전체 데이터는 그림 1과 같이  $G_1 \sim G_K$ 의  $K$  개의 그룹으로 분할되어 각 방송채널에 할당되는데, 한 데이터는 한 채널에만 할당될 수 있다.  $CH_i$ 는  $G_i$ 에 속한 데이터들을 라운드-로빈 방식으로 평형 방송하며, 각  $G_i$ 의 맨 마지막 데이터의 인덱스를  $ccut_i$ 로 표현한다.

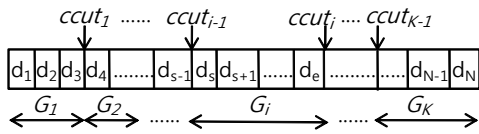


그림 1.  $N$  데이터들을  $K$  개의 그룹으로 분할하는 문제  
Fig. 1. Problem for partitioning  $N$  data items in  $K$  groups

- 4) 한 방송채널에서 하나의 데이터를 방송하는데 걸리는 논리적인 시간을 1 tick으로 정의하고, 응답시간 성능의 단위로서 tick을 사용한다.

## 2. 기본 개념 및 데이터 할당 알고리즘

단일 데이터 질의 환경의 데이터 할당 기법들은<sup>[3-7]</sup> 액세스 확률이 높은 데이터들이 배정되는 방송채널에는 적은 수의 데이터를 할당하여 이들의 방송빈도를 높이는 방향으로 평균 응답시간의 최적화를 시도했다. 따라서 액세스 확률이 낮은 데이터들은 한 채널에 많은 수가 할당되어 이들은 상대적으로 매우 낮은 빈도로 방송된다. 그런데 다중 데이터 질의에서는 가장 늦게 수신된 데이터에 의해 응답 시간이 결정된다. 따라서 질의 내의 어느 한 데이터가 방송 빈도가 낮아 액세스하는데 시간이 많이 걸린다면, 다른 데이터를 아무리 빨리 수신하더라도 응답시간 측면에서는 도움이 되지 않는다. 그러므로 다중 데이터 질의 환경에서는 액세스 확률이 방송빈도에 미치는 영향력을 적절하게 제어하지 않으면, 응답시간의 성능이 크게 저하될 가능성이 있다.

본 논문은 이와 같은 점을 고려하여 다음 두 가지의 직관적인 관찰을 데이터 할당 기법에 적용한다. 첫째, 액세스 확률이 높은 데이터는 질의에 포함될 가능성이 높기 때문에 액세스 확률이 높을수록 더 자주 방송되어야

한다. 둘째, 질의에 포함된 데이터 항목의 수가 많아짐에 따라 액세스 확률이 방송 빈도에 미치는 영향력이 작아져야 한다. 만일 극단적으로 전체 데이터들을 액세스하는 질의의 경우에는 데이터의 액세스 확률이 방송빈도에 전혀 영향을 미치지 않고 모든 데이터들을 동일 빈도로 방송하는 것이 가장 우수한 성능을 발휘할 것이다.

본 논문은 이와 같은 특징을 데이터 할당 정책에 반영하기 위하여 각 데이터의 액세스 확률  $p_i$ 의 이동 평균(Moving Average) 값을 식 (1)과 같이 계산하여 이를 방송 빈도를 결정하는데 사용한다. 이 수식에서  $Q$ 는 한 질의에 포함되는 평균 데이터 항목의 수를 의미하며,  $R$ 은  $Q$ 가 방송빈도에 미치는 영향력을 의미한다.

$$MA(p_i) = \left( \sum_{j=i}^{i+Q \times R - 1} p_j \right) / (Q \times R) \quad (1)$$

단일 데이터 환경에서는  $p_i$  값에 기초하여 방송빈도를 결정하는데, 본 논문은  $MA(p_i)$  값에 기초하여 방송빈도를 결정한다.  $Q$  값이 커지면 각 데이터 간에  $MA(p_i)$  값의 차이가 줄어들기 때문에 액세스 확률이 방송빈도에 미치는 영향력이 줄어들게 된다. 다중 데이터 질의에서  $p_i$ 와  $Q$ 가 주어진 상태에서  $R$  값을 어떻게 설정하는 것이 가장 최적의 응답시간 성능을 발휘하는지를 해석적으로 분석하는 연구는 매우 흥미 있는 과제이기는 하지만 본 연구에서는 이를 미래 과제로 남겨둔다. 본 논문에서는 다양한 환경에서 실험을 통하여 최적의  $R$  값을 찾아내 이를 적용할 것이다.

단일 데이터 질의에서는 각 데이터들의 방송빈도를  $\sqrt{p_i}$  값에 비례하게 설정할 때 가장 최적의 응답시간 성능을 발휘한다<sup>[12]</sup>. 본 논문은 다중 데이터 질의 환경에서는 각 데이터의 방송 빈도를  $\sqrt{MA(p_i)}$  값에 비례하게 설정할 때 우수한 성능을 발휘할 것이라는 휴리스틱을 적용하여 각 방송채널에 데이터를 할당한다. 제안된 기법은 이와 같이  $MA(p_i)$  값에 기초하여 데이터를 분할하여 방송채널에 할당하므로 DAMA(Data Allocation using Moving Average)라고 명명한다. DAMA의 데이터 분할 및 할당 알고리즘이 그림 2에 제시되어 있다.

이 알고리즘은 우선 후방향으로 각  $p_i$ 의 이동 평균  $MA(p_i)$ 를 계산한다. 그리고  $\sqrt{MA(p_i)}$ 를 구하면서  $\sum_{i=1}^N \sqrt{MA(p_i)}$  값을 계산하여  $SumSrMA$ 에 저장한다. 본 논문은 각 데이터들의 상대적인 방송빈도를  $\sqrt{MA(p_i)}$  값에 비례하게 설정할 때 우수한 성능을 발휘할 것이라

```

//precompute  $MA(p_i)$ 
for each data index  $i$  from 1 to  $N-Q \times R$ 
     $MA(p_i) = (\sum_{j=i}^{i+Q \times R-1} p_j) / (Q \times R)$ 
end for
for each data index  $i$  from  $N-Q \times R+1$  to  $N$ 
     $MA(p_i) = (\sum_{j=i}^N p_j) / (N-i+1)$ 
//precompute  $\sqrt{MA(p_i)}$  and  $\sum_{i=1}^N \sqrt{MA(p_i)}$ 
 $SumSrMA = 0$ 
for each data index  $i$  from 1 to  $N$ 
     $srma_i = \sqrt{MA(p_i)}$ 
     $SumSrMA = SumSrMA + srma_i$ 
end for
//determine  $ccut_0 \sim ccut_K$ 
 $AugSrMA = SumSrMA / K$ 
 $ccut_0 = 0, ccut_K = N$ 
①for each cut point  $i$  from 1 to  $K-1$ 
     $ChSumSrMA = 0$ 
    for each data index  $j$  from  $ccut_{i-1}+1$  to  $N$ 
         $ChSumSrMA = ChSumSrMA + srma_j$ 
        if ( $ChSumSrMA \geq AugSrMA$ )
             $ccut_i = j$ 
            break;
        end if
    end for
end for
for each cut point index  $i$  from 1 to  $K$ 
    allocate  $d_{ccut_{i-1}+1} \sim d_{ccut_i}$  into  $CH_i$ 
end for
    
```

그림 2. 데이터 분할 및 할당 알고리즘  
 Fig. 2. Data partitioning and allocating algorithm

는 휴리스틱을 사용하고 있기 때문에 이를 위해서는  $SumSrMA$ 를 각 방송채널에 균등하게 배분해야 한다. 그림 2의 ① 부위의 반복문에서는 이와 같은 개념으로 전체 데이터들의 분할점  $ccut_i(i=1 \sim K-1)$ 를 결정한다. 그림 1에서 보는 것과 같이  $ccut_i$ 는  $i$ 의 범위가  $1 \sim K-1$ 만이 의미가 있는데, 구현 편의를 위해  $ccut_0 = 0, ccut_K = N$ 으로 설정한다. 알고리즘 이해를 위하여  $ccut_1$ 을 결정하는 과정을 살펴보자.  $ChSumSrMA$ 는 각 방송채널에 할당되는 데이터들의  $\sum \sqrt{MA(p_i)}$ 를 저장하기 위한 임시 변수인 데 이를 우선 0으로 초기화한다. 그리고 나서  $d_{ccut_0+1}$ , 즉  $d_1$ 부터  $d_N$ 까지 순서대로 개념상  $G_1$ 에 포함시키면서  $ChSumSrMA$ 를 재계산한다.  $ChSumSrMA$  값이  $AugSrMA$ 와 같아지거나 이를 초과하게 되면 맨 마지막에 포함시킨 데이터가  $ccut_1$ 로 결정된다.  $ccut_1$ 이 결정되고 나면  $ChSumSrMA$ 를 다시 0으로 초기화하고  $d_{ccut_1+1} \sim d_N$ 을 순서대로  $G_2$ 에 포함시키면서  $ccut_2$ 를 결

정한다. 이와 같은 방식으로  $ccut_1$ 부터  $ccut_{K-1}$ 까지 순차적으로 구할 수 있다. 각  $ccut_i(i=0 \sim K)$ 가 결정되고 나면,  $CH_i$ 에는  $d_{ccut_{i-1}+1} \sim d_{ccut_i}$ 의 데이터들이 할당되어 라운드-로빈 방식으로 평형 방송된다.

## IV. 성능 평가 모델 및 성능 평가

이 장에서는 본 논문이 제안한 DAMA 기법의 성능을 평가하고 그 결과를 기술한다. 비교 대상으로는 FLAT<sup>[1]</sup>, DP<sup>[3]</sup>, 그리고, NODA<sup>[7]</sup> 기법을 선정하였다.

### 1. 성능평가 모델

본 논문의 성능평가 모델은 크게  $K$  개의 방송채널 서버와 모바일 클라이언트로 구성된다. 방송 시스템의 특성상 어떤 클라이언트의 성능은 다른 클라이언트들의 존재 여부에 영향을 받지 않기 때문에 본 논문은 클라이언트의 수를 하나로 고정한다. 클라이언트의 질의는  $Q$  개의 데이터로 구성되는데, 클라이언트는 현재 질의가 필요로 하는 모든 데이터를 수신한 후에 다음 질의를 제기한다. 클라이언트는 한 순간에는 한 방송채널만 수신할 수 있으며, 질의에 포함된 데이터들이 서로 다른 방송채널을 통하여 동시에 방송되면 가장 방송 빈도가 낮은 데이터를 우선 수신한다. 본 성능평가 모델은 MCC에서 개발한 CSIM<sup>[13]</sup>을 이용하여 구현하였으며, 본 논문에서 사용된 입력변수들은 표 1과 같다.

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        | 방송 채널의 수            | 10    | 1 ~ 20                           |
| $\theta$ | zipf 분포의 $\theta$ 값 | 0.95  | 0.0 ~ 1.0                        |
| Q        | 질의에 포함된 데이터 수       | 5, 10 | 1-15                             |
| R        | Q의 방송빈도 영향력         | 100   | 1-200                            |
| B        | NODA의 서버채널 수        | 10    | $\max(10, \lfloor 50/K \rfloor)$ |

서버가 방송하는 데이터들은 동일한 크기를 가진  $N$  개의 데이터로 구성되는데, 이들은 각 데이터 할당 기법에 맞추어 방송채널에 할당되어 주기적으로 방송된다. 그리고 각 데이터의 액세스 확률은 불균등한 액세스 형태를 모델링하기 위해 많이 사용되는 zipf 분포 모델을<sup>[1]</sup>

<sup>6)</sup> 따른다고 가정한다. zipf 분포 모델에서 각 데이터의 액세스 확률은 식 (2)로 표현되는데, 각 질의는 이 액세스 확률에 근거하여 질의에 포함시킬 데이터를 결정한다. 그리고 각 데이터 할당 기법들은  $p_i$  또는  $MA(p_i)$ 에 기초하여 데이터들을 분할하여 각 방송채널에 할당한다. 여기서  $\theta$  값은 각 데이터들의 액세스 확률이 어느 정도 편향되어 있는지를 나타내는 인수로써,  $\theta$  값이 0일 경우에는 각 데이터의 액세스 확률은 모두 동일하며  $\theta$  값이 커질수록 불균등한 액세스의 정도가 심화된다.

$$p_i = \frac{(1/i)^\theta}{\sum_{j=1}^N (1/j)^\theta} \quad (1 \leq i \leq N) \quad (2)$$

DAMA 기법은  $Q$ 가 방송빈도에 미치는 영향력을 의미하는  $R$  값을 60~150으로 설정할 때 우수한 성능을 보이기 때문에 이 값은 100으로 설정하고 실험을 실시했다. 단 단일 데이터 질의의 경우에는  $R$  값을 1로 설정하여, 액세스 확률에 기초하여 방송채널에 데이터를 할당한다. 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 개의 질의 결과는 무시하였다. 성능 결과에서 제시된 값은 5 개의 서로 다른 임의의 수를 사용하여 실시된 실험 결과의 평균값으로, 각 실험은 천만 개의 질의가 처리될 때까지 실시하였다.

## 2. 질의 크기에 따른 평균 응답시간의 성능

본 실험에서는 한 질의가 요청하는 데이터의 수를 의미하는 질의 크기를 1에서 15까지 변화시키면서 각 기법의 성능추이를 살펴보았다. 방송채널의 수는 10으로 설정했으며,  $\theta$  값은 데이터의 편향적인 액세스를 모델링하기 위한 대표 값으로 많이 사용되는 0.95로 설정하였다<sup>1)</sup>. <sup>6)</sup> 질의의 평균 응답시간 성능이 그림 3에 제시되어 있는데, DAMA는 대체로 다른 기법보다 우수한 성능을 보인다. 그래프 상으로 잘 구분이 되지 않지만, 질의 크기가 1~2 구간에서 NODA는 다른 기법보다 우수한 성능을 보인다. NODA 기법은 서브 채널의 개념을 사용하여 동일 채널에 할당된 데이터들도 그들의 인기도를 반영하여 다른 빈도로 방송하기 때문에  $Q$ 가 작을 경우에는 다른 기법보다 우수한 성능을 보일 수 있다.

그러나  $Q$ 가 커지게 되면 액세스 확률이 낮은 데이터도 질의에 포함될 가능성이 커지기 때문에 액세스 확률

이 방송빈도에 미치는 영향력이 작아져야 한다. DAMA 기법은  $Q \times R$  개 데이터의 액세스 확률의 이동평균 값에 기초하여 방송채널에 데이터를 할당하기 때문에  $Q$ 가 커짐에 따라 액세스 확률이 방송 빈도에 미치는 영향력을 감소시킬 수 있다. 따라서  $Q$ 가 3 이상이 되면 DAMA 기법은 NODA와 DP 기법보다 우수한 성능을 보이기 시작한다. 참고로 질의 크기가 10 이상이 되면, DAMA 기법은 NODA와 DP 기법에 비해 각각 13%, 17% 이상의 우수한 성능을 보이는데, 질의 크기가 커질수록 그 성능 차이는 점점 커진다.

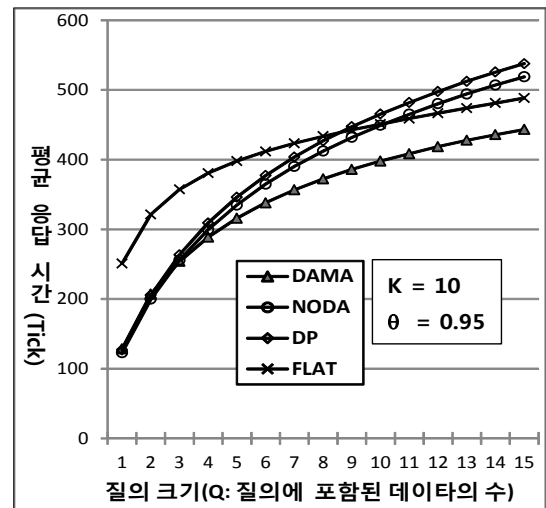


그림 3. 질의 크기에 따른 평균 응답시간  
Fig. 3. Query size vs. average response time

FLAT 기법은 액세스 확률에 관계없이 모든 데이터들을 동일 빈도로 방송한다. 따라서 이 기법은  $Q$ 가 작을 경우에는 다른 기법에 비해 매우 낮은 응답시간의 성능을 보인다. NODA와 DP 기법은 액세스 확률이 낮은 데이터의 방송빈도를 낮게 하는 정책을 사용함으로써 액세스 확률이 높은 데이터들을 빈번하게 방송할 수 있었다. 그러나 질의 크기가 커지면 액세스 확률이 낮은 데이터들도 질의에 포함될 가능성이 높아지기 때문에  $Q$ 가 증가하면 NODA와 DP 기법의 응답시간 성능은 다른 기법에 비해 급격하게 저하된다.  $Q$ 가 증가함에 따라 FLAT 기법과 DP 및 NODA 기법의 성능 차이는 줄어들다가,  $Q$ 가 11 이상이 되면 FLAT 기법이 NODA와 DP 기법보다 우수한 성능을 보이기 시작한다.

### 3. 채널 수의 변화에 따른 평균 응답시간의 성능

이 절에서는 방송채널의 수를 1~20으로 변화시키면서 각 기법의 성능 추이를 살펴보았다. 이 실험에서  $\theta$  값은 0.95로 설정하였다. 질의 크기를 5와 10으로 설정하고 실시한 실험의 평균 응답시간 성능이 각각 그림 4와 그림 5에 제시되어 있다.

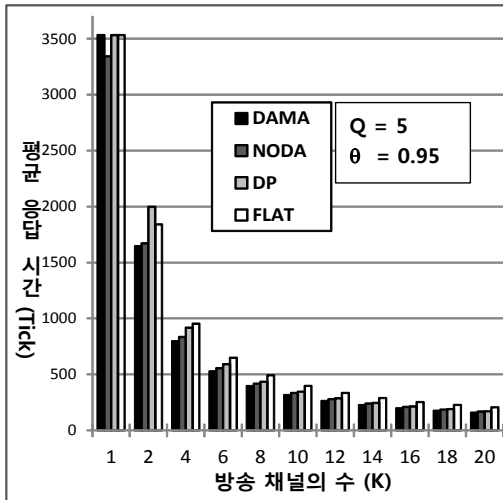


그림 4. 방송 채널의 수에 따른 평균 응답시간(Q=5)  
 Fig. 4. No. of channel vs. average response time(Q=5)

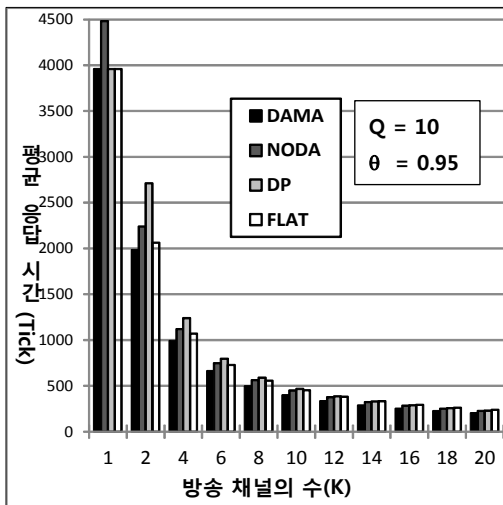


그림 5. 방송 채널의 수에 따른 평균 응답시간(Q=10)  
 Fig. 5. No. of channel vs. average response time(Q=10)

NODA를 제외한 다른 기법들은 모두 방송채널 별로 평형 방송을 실시하는 시스템 모델을 사용한다. 따라서  $K$ 가 1일 경우에는 DAMA, DP, 그리고 FLAT 기법은 모든 데이터들을 같은 빈도로 방송하게 되므로 이 기법들은 모두 동일한 성능을 보인다. 반면, NODA 기법은 동일 채널에 할당된 데이터들도 그들의 인기도를 반영하여 다른 빈도로 방송할 수 있기 때문에 타 기법들과는 다른 성능 결과를 보인다. 단일채널 환경에서 NODA 기법은 그림 4과 그림 5에서 보는 것과 같이  $Q$ 가 작을 경우에는 다른 기법보다 우수한 성능을 보이지만,  $Q$ 가 10으로 커지면 액세스 확률이 방송빈도에 미치는 영향력이 너무 크다는 단점으로 인해 더 나쁜 성능을 보이기도 한다.

그림 4와 그림 5를 살펴보면, 모든 기법에서 방송채널의 수가 증가함에 따라 질의 응답시간이 감소하는 것을 확인할 수 있다. 이는 방송채널의 수가 증가함에 따라, 각 채널에 할당되는 데이터의 개수가 작아져 데이터들을 더 빈번하게 방송할 수 있기 때문이다. 이 실험에서도 DAMA 기법은 단일 방송채널인 경우를 제외한 모든 채널 구간에서 다른 기법보다 우수한 응답시간의 성능을 보이는데, 질의 크기가 크면 대체로 더 큰 성능 차이를 보임을 알 수 있다. 이는 DAMA 기법이 질의 크기에 맞추어 데이터의 액세스 확률이 방송 빈도에 미치는 영향을 적절하게 제어할 수 있기 때문이다. 참고로 방송채널의 수를 4로 설정할 때의 성능을 살펴보면, 질의 크기가 5인 경우에 DAMA 기법은 NODA와 DP 기법에 비해 각각 5%, 15% 정도의 우수한 성능을 보이는데, 질의 크기가 10인 경우에는 13%, 25% 정도의 우수한 성능을 보인다.

### 4. $\theta$ 값의 변화에 따른 평균 응답시간의 성능

본 실험에서는  $\theta$  값의 변화에 따른 평균 응답시간의 성능추이를 살펴보았다.  $K$ 를 10으로 설정하고,  $\theta$  값을 변화시키면서 살펴본 성능 결과가 그림 6에 제시되어 있다.  $\theta$  값이 0이 되면, 각 데이터의 액세스 확률이 동일하게 되어 클라이언트가 액세스하는 데이터들의 분포는 균등 분포를 이루게 된다. 따라서 이 경우에는 각 기법이 어떤 데이터 할당 정책을 사용하더라도 동일한 성능을 보인다.  $\theta$ 가 증가함에 따라 각 데이터간의 액세스 확률의 차이는 점점 커지게 되며, 이에 따라 각 기법의 응답시간 성능은 향상됨을 알 수 있다.

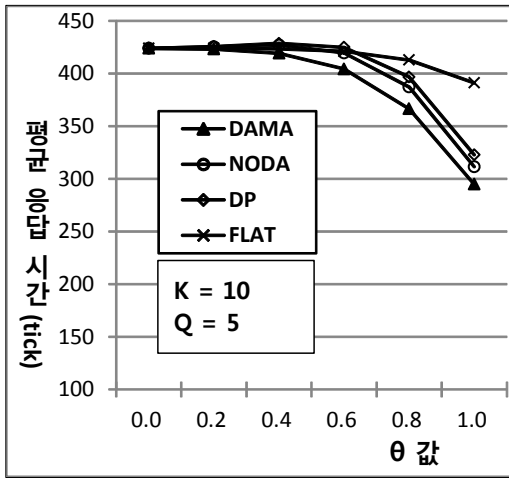


그림 6.  $\theta$  값의 변화에 따른 평균 응답시간  
 Fig. 6.  $\theta$  value vs. average response time

각 기법의 응답시간 성능이 개선되는 이유는 다음의 세 가지 원인에서 비롯된다. 첫째, FLAT을 제외한 다른 기법들은 액세스 확률의 차이를 고려하여, 정도의 차이는 있지만 액세스 확률이 높은 데이터일수록 더 자주 방송을 하기 때문이다. 둘째,  $\theta$ 가 커지면 질의가 필요로 하는 데이터의 동시 방송 가능성이 줄어들기 때문이다. 만일 질의의 데이터들이 동시에 방송되지 않는다면, 질의의 최악 응답시간은 한 방송주기를 초과하지는 않는다. 그렇지만 두 데이터가 동시에 방송되면 이 중 하나를 수신하고, 다른 데이터는 다음 방송 주기에서 수신해야 하므로 응답시간이 길어질 수 있다.  $\theta$  값이 커지면, 질의에 포함된 데이터들의 대부분은 액세스 확률이 높은 데이터들이 할당된  $CH_i$ 가 방송하게 될 것이다. 동일 채널에 할당된 데이터들은 동시에 방송될 수 없기 때문에  $\theta$  값이 커지면 질의에 포함된 데이터의 동시 방송 가능성이 줄어들게 된다. 참고로  $\theta$ 가 0과 1.0 일 경우에 각각 18만 개, 10만 개 정도의 데이터가 동시에 방송됨을 관찰할 수 있었다. 셋째,  $\theta$ 가 증가함에 따라 질의가 방송채널의 방송 순서상 인접한 데이터들로 구성될 가능성이 높아지기 때문이다. 참고로  $\theta$ 가 1.0이 되면  $d_1 \sim d_{30}$ 의 액세스 확률이 합이 44% 정도가 된다. 따라서 이들 데이터만으로 질의가 구성될 가능성이 높아지게 된다. 모든 기법에서 방송채널의 데이터는 그림 1과 같이 데이터의 순서대로 할당되어 있기 때문에 인접한 데이터들로 질의가 구성되면 응답시간 성능이 향상된다. 이에 대한 이해를 돕기 위해 어떤 채널에

$d_1 \sim d_{400}$ 의 데이터가 할당되어 있다고 가정하자. 이 경우에  $\{d_1, d_2, d_3, d_4\}$ 를 액세스하는 질의의 평균 응답시간은 대략 방송주기의 반인 200 tick이 될 것이다. 그러나  $\{d_1, d_{101}, d_{201}, d_{301}\}$ 을 액세스 하는 질의는 350 tick 정도의 평균 응답시간을 보일 것이다.

이런 세 가지 이유의 조합으로 인하여  $\theta$ 가 증가함에 따라 각 기법의 응답시간 성능은 향상되는데, DAMA 기법은 액세스 확률의 차이와 질의 크기를 적절하게 고려하여 방송채널에 데이터를 배정하기 때문에 다른 기법보다 우수한 성능을 보인다. NODA와 DP 기법의 응답시간 성능도 개선되지만, DAMA 기법만큼 급격하게 향상되지는 않는다. 이는 NODA와 DP 기법에서는 액세스 확률이 방송 빈도에 미치는 영향력이 너무 크기 때문이다. FLAT 기법은 모든 데이터들을 동일 빈도로 방송하기 때문에  $\theta$  값에 관계없이 일정한 성능을 보일 것 같지만 실제로는 그렇지 않다. 그 이유는 위에서 설명한 두 번째와 세 번째의 원인에서 기인되는 것이다.

## V. 결론

본 논문은 방송 방식의 데이터 전달 기법을 채택한 다중 방송채널 환경의 정보시스템에서 다수의 데이터를 검색하는 다중 데이터 질의를 효율적으로 처리하기 위한 데이터 할당 기법에 대한 연구를 수행했다. 본 논문은 다중 데이터 질의의 응답시간을 최소화시키는 것을 목표로 하여 DAMA라 명명된 새로운 데이터 할당 기법을 제안했다. DAMA 기법은 액세스 확률이 높은 데이터일수록 질의에 포함될 가능성이 높아 방송빈도가 빈번해야 하며, 질의에 포함된 데이터의 수가 많아짐에 따라 액세스 확률이 방송빈도에 미치는 영향력이 작아져야 한다는 점을 고려하여 액세스 확률의 이동평균 값에 기초하여 데이터를 분할하여 방송채널에 할당한다. 본 논문은 다중 데이터 질의를 위한 시뮬레이션 모델을 설계 구현하여 성능 평가를 실시하였다. 본 논문이 제안한 기법은 질의 크기에 따라 액세스 확률이 방송빈도에 미치는 영향력을 적절하게 제어할 수 있기 때문에 대부분의 환경에서 다른 기법과 비교하여 우수한 평균 응답시간의 성능을 보였다. 본 논문은 각 데이터의 액세스 확률과 질의 크기만이 주어진 환경에서 다중 데이터 질의의 응답시간을 개선하기 위한 연구를 진행하였다. 본 논문의 미래 연구로서는



질의 패턴 및 질의의 요청 빈도를 파악할 수 있는 환경에서 다중 데이터 질의를 처리하기 위한 연구를 진행할 것이다.

## References

- [1] S. Acharya, "Broadcast Disks: Dissemination-based Data Management for Asymmetric Communication Environments," Ph.D. thesis, Brown University, 1998.
- [2] M. Franklin and S. Zdonik, "A Framework for Scalable Dissemination-Based Systems," In the Int. Conf. on Object-Oriented Programming Languages, Systems, and Applications, pp. 94-105, Atlanta, GA, October 1997.
- [3] 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.
- [4] 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.
- [5] S. Wang and H.L. Chen, "Near-optimal data allocation over multiple broadcast channels," Computer communications, Vol. 29, No. 9, pp. 1341-1349, Elsevier Science, 2006.
- [6] H. M. Kwon, "TLDP: A New Broadcast Scheduling Scheme for Multiple Broadcast-Channel Environments," The Journal of the Institute of Internet, Broadcasting and Communication(JIIBC), Vol. 11, No. 2, pp. 63-72, 2011.
- [7] H. M. Kwon, "A Near Optimal Data Allocation Scheme for Multiple Broadcast-Channel Environments," The Journal of the Institute of Internet, Broadcasting and Communication(JIIBC), Vol. 12, No. 1, pp. 17-27, 2012.
- [8] Y. D. Chung and M. H. Kim, "Effective Data Placement for Wireless Broadcast," Distributed and Parallel Databases, Vol 9, No. 2, 2001
- [9] G Lee and SC Lo, "Broadcast data allocation for efficient access of multiple data items in mobile environments," Mobile Networks and Applications, Vol. 8, No. 4, pp. 365-375, 2003.
- [10] 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.
- [11] H.P. Hung, J.W. Huang, J.L. Huang, and M.S. Chen, "Scheduling dependent items in data broadcasting environments," ACM SAC 2006.
- [12] 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.
- [13] 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 2013학년도 세명대학교 교내학술연구비 지원에 의해 수행된 연구임