

사용자 캐쉬를 사용한 데이터 방송 시스템에서의 사용자 대기시간 분석

(Analysis of Waiting Time in a Data Broadcast System using User Cache)

김 종 덕 * 김 종 권 **

(Jong-Deok Kim) (Chong-Kwon Kim)

요 약 데이터 방송은 무선 통신 환경에서 많은 사용자들에게 정보를 전달하는 데 있어 효과적인 방안으로 주목받고 있다. 데이터 방송 시스템에서의 중요 문제 중 하나는 사용자 기다림을 최소화하는 것이다. 사용자 대기시간을 줄이는 데는 두 가지 접근법이 있다. 하나는 대기시간을 최소화하는 방송 스케줄의 설계이고, 다른 하나는 사용자 시스템의 지역 메모리를 캐쉬로 이용하여 사용자 요구에 즉각 응답하는 것이다. 이 두 가지 접근법은 동시에 적용되어 전체 성능을 보다 좋게 할 수 있다. 효율적인 방송 시스템의 설계 및 구현을 위해서는 방송 스케줄, 캐쉬 크기, 캐싱 정책 등 여러 가지 요인을 종합적으로 고려하는 것이 필요하다. 본 논문에서는 WWW과 같이 정보 항목들간에 구조적 연관 관계가 존재하는 시스템에서, 앞서 기술한 여러 가지 요인을 함께 고려하여 방송 시스템의 성능을 분석하였다. 분석적 기법을 이용하여 채택한 방송 스케줄, 캐쉬 크기, 캐싱 정책의 조합으로 얻을 수 있는 시스템 성능의 하한을 구하였다. 이 결과를 이용하면 여러 가지 시스템 구현 방안간의 성능 비교가 가능하며 이를 통해 효과적인 방송 시스템의 설계 및 구현이 가능할 것이다. 우리는 분석의 타당성 검증을 위해 시뮬레이션을 수행하였다.

Abstract Data broadcast has been considered a promising way of information dissemination to a massive number of users in a wireless communication environment. Reducing user waiting time is a major problem in developing a data broadcast system. There are two approaches for this problem: One is to design a broadcast schedule at the server side which reduces the mean response-time, and the other is to utilize a local cache at the user side which may respond a user request instantly. Though these two approaches were addressed separately in literature, they may be taken jointly for better performance. The performance of system with joint approach depends on several factors such as broadcast schedule, cache size, cache management strategy, etc. In this paper, we analyze user waiting time in a data broadcast system with joint approach in which information items are structurally related with each other as in WWW. Based on the worst case assumption, we derive a lower bound of the system performance for given broadcast schedule, cache size, and cache management strategy. This result will be of help for designing and developing a data broadcast system. We support our analysis by carrying out extensive simulation on some interesting proposed broadcast schedules and cache management strategies.

1. 서 론

데이터 방송은 무선 통신 환경에서 많은 사용자들에

게 정보를 전달하는 데 있어 효과적인 방안으로 주목받고 있다. 상대적으로 높은 비용이 드는 사용자로부터 서버로의 통신을 줄일 수 있으며, 서버의 부하가 서비스 대상 사용자 수에 독립적이므로 확장성이 뛰어나기 때문이다. 실제 이를 채택한 현실 응용 시스템들도 여럿 있는데 BellCore의 Datacycle Project, Boston Community Information System, Intel의 InterCast 등이 있다[4]. 이러한 시스템들을 통해 전달되는 정보로는

* 비 회 원 : 서울대학교 컴퓨터공학부
kimjd@popeye.snu.ac.kr

** 종신회원 : 서울대학교 컴퓨터공학부 교수
ckim@popeye.snu.ac.kr

논문접수 : 2000년 4월 18일

심사완료 : 2000년 10월 2일

뉴스, 날씨 정보, 도로 정보, 항공기 및 열차 스케줄, 주식 시세 등이 있다. 무선 통신 환경에서 동영상도 포함하는 멀티미디어 정보의 방송이 필요성이 날로 증대되고는 있지만 대역폭 등의 문제로 가까운 미래까지는 앞서 살펴본 데이터 정보의 방송이 보다 현실적이고 유용한 응용의 위치를 지킬 것이다. 그리고 멀티미디어 정보는 일반 데이터 정보와는 구별되는 정보 모형 및 특성을 가지며 일반 데이터 방송과는 다른 방식을 통해 처리되어야 할 필요성이 있다. 본 논문에서 무선 통신 환경에서의 멀티미디어 정보의 방송 문제는 다루고 있지 않다.

종종 무선 통신 환경에서는, 사용자로부터 정보 제공 서버로의 명시적인 정보 요구 전달이 가능하지 않거나 또는 적합하지 않을 수 있다. 따라서 단방향 푸쉬 기반 데이터 방송 시스템이 주요 관심 대상이 된다. 이 시스템에서 서버는 사용자로부터의 아무런 되먹임 없이 사용자들이 필요로 하는 정보 항목들을 주기적이고 지속적으로 방송한다. 특정 정보 항목을 얻기 위해서 사용자는 서버가 해당 정보 항목을 방송할 때까지 기다려야만 한다.

방송 시스템에서 발생하는 이러한 필연적인 사용자 기다림을 줄이는 것은 방송 시스템 구축에 있어 해결해야 할 핵심 문제가 되며 기본적으로 두 가지 접근법이 있다. 하나는 서버 측 방안으로 평균 대기시간을 최소화하는 방송 스케줄을 설계하는 것이다. 다른 하나는 사용자 측 방안으로 사용자 시스템의 지역 메모리를 캐쉬로 이용하여 대기시간을 줄이는 것이다. 즉 방송되어 오는 정보 항목들 중 적절한 것을 선택, 미리 수신하여 캐쉬에 저장해 두었다가 해당 항목을 사용자가 요구할 때 즉각 응답하는 것이다.

방송 시스템에서의 대기시간을 줄이기 위해 그 동안 많은 연구들이 수행되었다. 특히 평균 대기시간을 최소화하는 방송 스케줄을 얻기 위한 연구가 활발하게 수행되었다. Vaidya 등은 방송 스케줄과 사용자 대기시간의 평균 및 분산과의 연관 관계를 분석하고, 각각을 최소화할 수 있는 조건식을 유도하였다. 일반적으로 대기시간 평균의 최소화와 분산의 최소화는 동시에 이룰 수 없는 목표이다. 이들은 평균과 분산을 서로 타협하여 두 가지 모두를 적정하게 줄일 수 있는 알고리즘을 제시하였다[1,2]. 또한 사용자 캐싱 정책과 관련한 연구 결과 우선 순위를 평가할 때 항목의 요청 확률 외에 재방송까지 남은 시간을 고려할 경우 성능을 향상시킬 수 있음을 발견하였다[5,6,7]. Su와 Tassiulas는 서버의 방송 스케줄과 사용자 캐싱 정책을 적절히 조합하여 평균 대

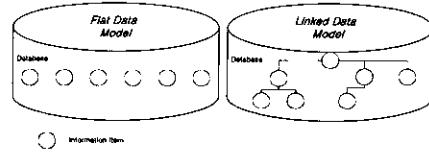


그림 1 Flat Data Model과 Linked Data Model

기 시간을 최소화하는 방안에 대한 연구를 수행하였다[4].

이러한 연구들은 우리가 "Flat Data Model"(편의상 이후로는 FDM이라고 줄여 쓴다)이라 이름 지은, 정보 항목들간에 아무런 구조적 연관 관계를 가정하지 않는 정보 모형을 기반으로 수행되었다. 그러나 실제 많은 정보 시스템의 정보 데이터베이스는 계층적 연관 구조를 가진다. 각 정보 항목은 다른 정보 항목과 구조적으로 연관되며 이에 대한 정보를 명시적으로 포함한다. 하이퍼 링크 정보를 명시적으로 포함하는 WWW의 HTML 문서는 대표적 예라고 할 수 있다. 우리는 이러한 형태의 정보 모형을 "Linked Data Model"(편의상 이후로는 LDM이라고 줄여 쓴다)이라고 이름 짓는다(그림 1 참조). Ammar도 텔레텍스트 시스템에서 개별 사용자 입장에서의 대기시간 분석을 위해 유사한 정보 모형을 이용하였다[3]. 하이퍼 링크 개념이 WWW의 대표적 성공 요인의 하나로 간주되는 것에 비추어 볼 때, 고려 중인 무선 환경에서의 방송 시스템에서 이와 유사한 개념의 채택이 유망한 것으로 보인다. LDM을 이용할 때 얻을 수 있는 가장 큰 장점의 하나로는 사용자 요구의 집약성을 쉽게 찾아 낼 수 있다는 점이다. 이에 따라 사용자 요구에 대한 효과적 예측이 가능해지고 사용자 캐쉬의 성능을 크게 향상시킬 수 있을 것이다.

그런데 LDM 기반 방송 시스템의 성능은 방송 스케줄, 캐쉬 크기, 캐싱 정책 등 여러 가지 요인의 조합에 의해 결정된다. 방송 시스템의 효과적 설계 및 구현을 위해서는 여러 요인들을 함께 고려한 이론적 성능 평가 도구가 필요하다. 우리는 분석적 기법을 이용하여 채택한 방송 스케줄, 캐쉬 크기, 캐싱 정책의 조합으로 얻을 수 있는 시스템 성능의 하한을, 사용자 대기 시간의 평균 및 분산 값을 척도로 구하였다. 이를 이용하면 여러 가지 구현 방안간의 성능 비교가 가능하다. 우리는 분석의 타당성 검증을 위해 제안된 주요 방송 스케줄 방안 및 사용자 캐싱 정책에 대하여 심도 있는 시뮬레이션을 수행하였다. 시뮬레이션 결과와 분석 결과의 비교를 통해 분석이 타당하며 효과적인 성능 평가 도구로 이용할 수 있음을 확인하였다.

논문의 나머지 부분은 다음과 같이 구성된다. 2장에

서는 분석에 이용한 LDM 및 사용자 캐쉬를 가지는 단 방향 푸쉬 기반 방송 시스템 모형을 소개한다. 3장에서는 방송 스케줄, 캐싱 정책, 사용자 모형과 사용자 대기 시간과의 연관 관계를 분석하고 식으로 도출한다. 4장에서는 분석의 타당성 검증을 위해 수행한 시뮬레이션에 대해 설명하고 그 결과를 비교, 분석한다. 마지막 5장에서 결론 및 앞으로의 연구 방향에 대하여 기술한다.

2. 시스템 모형

그림 2는 이 논문에서 가정하는 방송 시스템의 개략적 구조를 보여준다. 이 방송 시스템 모형은 3개의 부분으로 구성되는데, 정보 모형, 사용자 모형, 그리고 서버 방송 스케줄이다. 이 장에서는 시스템을 구성하는 이 세 가지 모형을 차례로 설명한다.

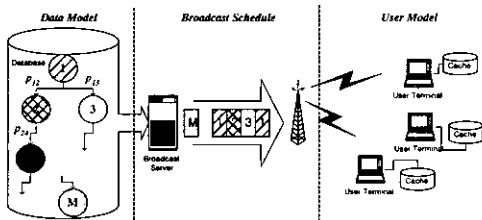


그림 2 Linked Data Model과 사용자 캐쉬를 가진 데이터 방송 시스템 모형

2.1 정보 모형

서버에서 제공하는 전체 정보 항목의 수를 M 이라 하자. 각 정보 항목의 길이(크기)는 동일하다고 가정하고 각 항목에게 1부터 M 까지 번호를 부여한다. LDM에서 각 정보 항목은 연결 항목들의 목록을 가진다. $R_u(n)$ 을 사용자 u 의 n 번째 요청 항목이라 하자. FDM에서 $R_u(n+1)$ 는 서버에 있는 임의의 정보 항목일 수 있지만, LDM에서는 $R_u(n)$ 의 연결 항목들 중의 하나로 제한된다. FDM의 경우 어떤 항목을 요청할 확률이 바로 전 요청 항목에 독립적이지만, LDM은 그것에 종속적이다.

두 정보 모형의 특징을 확률적으로 살펴보도록 하자. FDM에서는, 임의의 사용자 요구가 항목 i 를 요청할 확률을 항목 i 의 전역 요청 확률이라고 정의한다. 모든 i 에 대하여 전역 요청 확률이 시간에 무관하게 같다고 가정하고 그 확률을 π_i 이에 따른 확률 벡터를 $\pi = (\pi_i)$ 로 표시한다. $\sum_i \pi_i = 1$ 이 성립한다. LDM에서는 전역 요청 확률이 별도로 정의되지 않는다. 대신 어떤 사용자의 바로 전 요청 항목이 i 일 때, 그 사용자의 다

음 요청 항목이 j 일 확률을 항목 i 로부터 항목 j 로의 전이 확률이라 정의한다. 가능한 모든 (i, j) 쌍에 대하여 전이 확률이 시간에 무관하게 같다고 가정하고 그 확률을 p_{ij} 이에 따른 전이 확률 행렬을 $P = (p_{ij})$ 로 표시한다. $\sum_j p_{ij} = 1$ 가 성립한다. LDM에서는 추가로 가능한 모든 (i, j) 쌍에 대하여 i 로부터 j 로 가는 경로가 존재한다고 가정한다. 이로부터 LDM은 이산 마코프 연쇄가 된다.

2.2 사용자 모형

2.2.1 사용자 동작 모형

그림 3에 표시된 동일한 개별 사용자 동작 모형을 가지는 많은 사용자가 시스템에 있다고 가정한다. 사용자 수가 충분히 많을 경우, 합쳐진 사용자 요구는 일정한 발생 빈도를 가지는 포아송 프로세스를 따른다고 가정할 수 있다[1].

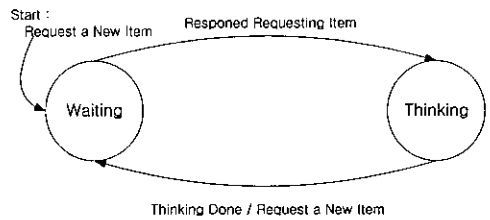


그림 3 개별 사용자 동작 모형

개별 사용자 동작 모형에서 사용자는 Waiting과 Thinking이라는 두 가지 상태에 있을 수 있다. Waiting은 어떤 항목을 요청하고 그 항목이 응답될 때까지 기다리는 것이고, Thinking은 해당 항목이 응답된 후 해당 항목에 대하여 사용자가 얼마간의 시간 동안 Thinking을 하는 상태이다. Thinking이 끝나고 사용자는 새로운 항목을 요청하며 다시 Waiting 상태로 돌아간다. 우리는 항목 i 에 대한 Thinking 시간을 λ_i 를 인자로 하는 지수 분포를 따른다고 가정한다. 개별 사용자 동작 모형의 타당성은 Ammar에 의해 검증된 바 있다[3].

2.2.2 사용자 캐쉬 모형

사용자는 K 개의 정보 항목을 저장할 수 있는 캐쉬를 가진다. 서버에 의해 어떤 항목이 방송될 때마다 사용자는 새로 방송되는 항목과 현재 캐쉬에 저장되어 있는 항목을 캐싱 정책에 따라 교체할 수 있다. 시각 t 에서 캐쉬에 존재하는 K 개의 정보 항목 집합을 C_t 로 표현하자. 시각 t 에서 발생한 항목 i 에 대한 요구는 $i \in C_t$ 이면 즉각 응답될 수 있지만 $i \notin C_t$ 이면 t 이후 이루어지

는 항목 i 에 대한 첫 번째 방송에 의해 응답된다.

대기시간 분석을 위해서는 시각 t 에서 특정 항목 i 의 캐쉬 적중 여부를 알 수 있어야 한다. 시각 t 이전에 있었던 항목 i 의 최종 방송 시각을 τ_i 라고 하자. $i \in C_t$ 라면 항목 i 는 C_t 의 항목들 및 (τ_i, t) 구간에서 방송된 항목들과 비교해 채택한 캐싱 정책에서 K 이내의 우선 순위를 가져야 한다. 그런데 이론적 분석에서 C_t 의 항목들 및 (τ_i, t) 구간에서 방송되는 항목들을 파악하는 일은 아주 어렵다. 분석을 어렵게 하는 또 다른 요인은 LDM의 특성을 반영하는 캐싱 정책의 경우 현재 항목이 변함에 따라 각 항목별 우선 순위가 변한다는 점이다. 현재 항목의 연결 항목들과 이전 항목의 연결 항목들이 상당 부분 다르기 때문에 항목별 우선 순위가 새로운 요청 전, 후로 크게 달라진다. 따라서 캐쉬가 크지 않을 경우 새로운 요청 이후, 캐쉬에 있던 기존 항목들은 대부분 교체되어 사라질 수 있다.

우리는 이러한 특성을 고려하여 분석의 편의를 위해 $i \in C_t$ 임을 확인할 수 있는 경우만 캐쉬 적중에 의한 응답이 이루어질 수 있다고 가정하는 최악 분석 기법을 적용하였다. 이를 위해 $i \in C_t$ 임을 확인할 수 있는 다음 두 가지 전체 조건을 제시한다. 첫 번째는 최종 방송 시각에 대한 조건이다. 앞서 지적하였듯이 새로운 요청이 발생하면 이전의 캐쉬 정보는 교체될 가능성이 아주 높으므로, (τ_i, t) 구간 사이에 새로운 요청이 발생하지 않는, 즉 τ_i 가 이전 항목의 Thinking 구간에 속하는 것을 $i \in C_t$ 이기 위한 첫 번째 전체 조건으로 한다. 두 번째는 전체 경쟁 조건으로 최종 방송 시각 조건이 성립한 이후 적용되는 것이다. C_t 의 항목들 및 (τ_i, t) 구간에서 방송되는 항목들은 항목 i 의 캐싱 경쟁 대상이라고 할 수 있으며 최악의 경우 이들은 전체 항목 집합일 수 있다. 전체 경쟁 조건은 항목 i 가 전체 항목 집합과 비교하여 상위 K 이내의 우선 순위를 가지는 것을 $i \in C_t$ 이기 위한 두 번째 전체 조건으로 하는 것이다.

최종 방송 시각 조건은 캐쉬 크기 및 캐싱 정책에 독립적이지만 전체 경쟁 조건은 이들과 연관된다. 유의하여야 할 사실은 실제로는 위 두 가지 조건을 만족하지 않아도 $i \in C_t$ 일 수 있으며, 따라서 위 조건을 통해 분석한 캐쉬의 성능은 실제 캐쉬 성능의 하한이 된다는 점이다.

2.3 방송 스케줄

이론적 분석을 위해 기존 방송 스케줄 연구 결과에서 채택한 중요 개념은 등간격 속성이다. 등간격 속성이란

각 항목의 방송간격이 등간격이라는 것을 의미한다. 다시 말하면 항목 i 에 대한 k 번째 방송 시작 시점을 t_k 라고 할 때, $t_{k+1} - t_k$ 의 값이 k 의 값에 독립적으로 일정한 상수 값 s_i 를 가진다는 것이다. 각 항목별 방송 비율이 결정되어 있을 때, 등간격 속성을 가지는 방송 스케줄이 그렇지 않은 것에 비해 보다 나은 성능을 가진다는 점이 기존 연구를 통해 밝혀졌다[1,2]. 그런데 등간격 속성은 FDM에서 얻어진 결과이며, 알려진 많은 방송 스케줄 방안이 이를 바탕으로 구축되었다.

앞서 가정한 충분히 많은 사용자들이 있다는 사용자 모형은, 서버의 관점에서, 임의의 사용자 요구가 항목 i 를 요청할 확률인 전역 요청 확률 π_i 와, 이산 마코프 연쇄인 LDM에서의 항목 i 에 대한 균형 상태 확률 $\bar{\pi}_i$ 가 같다고 가정할 수 있게 한다. 균형 상태 확률 벡터를 $\bar{\pi} = (\bar{\pi}_i)$ 라면 $\bar{\pi} = \bar{\pi}P$ 이 성립하고, 이로부터 각 항목의 전역 요청 확률(=균형 상태 확률)를 구할 수 있다. 즉 서버의 관점에서 관찰하면 LDM으로부터 기존의 FDM과 동일한 상황을 도출할 수 있는 것이다. 따라서 FDM을 바탕으로 하는 기존 방송 스케줄 방안들을 LDM에서도 그대로 사용할 수 있다. 앞으로 수행할 분석에서는 등간격 속성을 가지는 방송 스케줄만을 고려한다.

3. 분석

임의의 사용자 요구에서 사용자의 바로 전 요청 항목이 i 이고 새로 요청한 항목이 j 일 때의 조건부 평균 대기 시간을 μ_{ij} 라고 하자. i, j 는 1부터 M 까지의 값을 가지는 이산 확률 변수라 할 수 있다. 요청 항목이 j 일 확률은 이전 항목 i 에 종속되며, 앞에서 전이 확률 p_{ij} 로 정의하였다. 이전 항목이 i 일 확률은 항목 i 의 균형 상태 확률 $\bar{\pi}_i (= \pi_i)$ 라 할 수 있다. 따라서 전체 평균 대기 시간 μ 는 다음과 같다.

$$\mu = \sum_{i=1}^M \pi_i \cdot \left(\sum_{j=1}^M p_{ij} \cdot \mu_{ij} \right) \quad (1)$$

이제 조건부 평균 대기 시간 μ_{ij} 를 구한다. 시각 T 에서 발생한 요청 항목이 j 이고 해당 사용자의 바로 전 요청 항목이 i 인 임의의 사용자 요구에 대한 대기시간을 나타내는 확률 변수를 Ω_{ij} 라 하자. 시각 T 로부터 항목 j 가 T 이후 처음 방송되는 시점까지의 시간 간격을 나타내는 확률 변수를 L 이라 하자. 확률 변수 L 은 항목 j 의 캐쉬 적중 여부에 독립적이다. 항목 j 의 캐쉬 적중 여부를 나타내는 지시 함수 $\mathcal{W}(i, j, T)$ 을 (2)와 같이 정의한다.

$$\varphi(i, j, T) = \begin{cases} 0 & \text{if } j \in C_T \\ 1 & \text{if } j \notin C_T \end{cases} \quad (2)$$

확률 변수 Ω_{ij} 와 조건부 평균 대기 시간 μ_{ij} 는 다음을 만족한다.

$$\Omega_{ij} = \varphi(i, j, T) \cdot L, \quad \mu_{ij} = E[\Omega_{ij}].$$

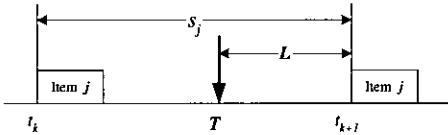


그림 4 T, L, t_k 간의 관계

확률 변수 L과 관련하여 항목 j의 방송 스케줄만이 관심의 대상이 된다. 요청 시각 T까지 항목 j가 k번 방송되었다면, 시각 T는 t_k 와 t_{k+1} 사이에 값을 가지며 $L = t_{k+1} - T$ 가 성립한다 (그림 4 참조). 사용자 모형에 의해 합쳐진 사용자 요구는 포아송 프로세스를 따르므로 확률 변수 T는 $(t_k, t_{k+1}]$ 범위에서 항등 분포한다고 할 수 있다. 따라서 확률 변수 L 역시 $(0, t_{k+1} - t_k]$ 범위에서 항등 분포한다. 등간격 속성에 의해 $t_{k+1} - t_k$ 는 k에 관계없이 상수 s_j 와 같다고 하였으므로 확률 변수 L은 $(0, s_j]$ 범위에서 항등 분포한다.

T의 시점에서 볼 때, 항목 j의 바로 전 방송 시점 t_k 는 등간격 속성에 의해 $[T - s_j, T)$ 범위에 있다고 할 수 있다. 확률 변수 Y를 $T - s_j$ 부터 t_k 까지의 시간, 즉 $Y = t_k - T + s_j$ 라 하자. $t_{k+1} = t_k + s_j$ 가 성립하므로 $L = Y$ 가 성립한다. 이로부터 다음을 얻을 수 있다.

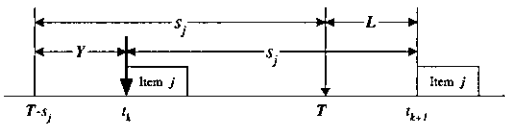


그림 5 T를 고정하고 관찰한 Y, L, t_k 간의 관계

$$P(Y=y) = \begin{cases} \frac{1}{s_j} & , 0 < y \leq s_j \\ 0 & , \text{otherwise} \end{cases} \quad (3)$$

$$\Omega_{ij} = \varphi(i, j, T) \cdot Y \quad (4)$$

이제 캐쉬의 영향을 고려하자¹⁾. 앞서 캐쉬 모형에서 지적하였듯이 최종 방송 시각 및 전체 경쟁 조건을 만족시켜 캐쉬를 확신할 수 있는 경우만 캐쉬 적용이 된다고 가정하고 분석한다. 이전 항목 i에 대한 Thinking

시간 $X_i = x_i$ 가 주어졌을 때, 새로 요청한 항목 j에 대한 방송 간격 s_j 가 x_i 보다 작다면, 항목 j는 등간격 속성에 의해 최종 방송 시각 조건을 언제나 만족한다. 그러나 방송 간격 s_j 가 x_i 보다 큰 경우, 최종 방송 시각 조건 만족을 보장할 수 없다. 이와 같이 최종 방송 시각 조건은 방송 간격 및 Thinking 시간과 밀접하게 관련되며 이를 고려하기 위한 지시 함수를 (5)와 같이 정의한다.

$$\delta(x_i, s_j) = \begin{cases} 0 & x_i > s_j \\ 1 & x_i \leq s_j \end{cases} \quad (5)$$

최종 방송 시간 조건을 만족하더라도 전체 경쟁 조건을 만족하지 못하면 캐쉬 적용을 확신할 수 없다. 전체 경쟁 조건을 위한 지시 함수 $\Delta_m(i, j, K)$ 를 정의한다. $\Delta_m(i, j, K)$ 는 캐쉬의 크기가 K이고 캐쉬 정책이 m인 캐쉬에서, 현재 항목이 i일 때 전체 경쟁 조건에서 항목 j의 캐쉬 여부를 나타내는 함수로 1이면 캐쉬가 되지 않음을 나타내고 0이면 캐쉬가 됨을 나타낸다.

지금까지의 관찰 결과를 바탕으로 주어진 항목 i에 대한 Thinking 시간이 x_i 로 주어졌을 때 Ω_{ij} 의 조건부 기대값을 $x_i > s_j$, $x_i \leq s_j$ 두 가지 경우에 대해 각각 구한다.

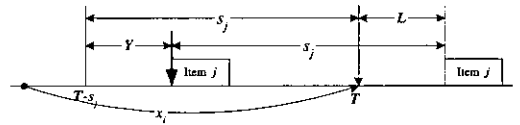


그림 6 $\delta(x_i, s_j) = 0$ 경우

• Case 1. $\delta(x_i, s_j) = 0$, (그림 6 참조)

가능한 모든 Y 값에 ($0 < y \leq s_j$) 대하여, 항목 j는 최종 방송 시각 조건을 만족한다. 항목 j의 캐쉬 여부는 $\Delta_m(i, j, K)$ 에 의해 결정되며, 따라서 $\varphi(i, j, T) = \Delta_m(i, j, K)$ 이 성립한다. (3), (4)로부터 다음을 유도할 수 있다.

$$E(\Omega_{ij} | X_i = x_i \wedge \delta(x_i, s_j) = 0) = \Delta_m(i, j, K) \cdot \int_0^{s_j} y \cdot \frac{1}{s_j} dy = \Delta_m(i, j, K) \cdot \frac{s_j}{2} \quad (6)$$

1) 캐쉬가 없을 경우 함수 $\varphi(i, j, T)$ 는 항상 1(=최종 실패)의 값을 가지며 (4)에서 $\Omega_{ij} = Y$ 가 된다. 등간격 속성 가정 아래 사용자 캐쉬가 없을 경우의 대기시간 평균 및 분산을 구한 기존 연구 결과와 우리의 분석 결과가 서로 일치함을 부록 1에서 보였다.

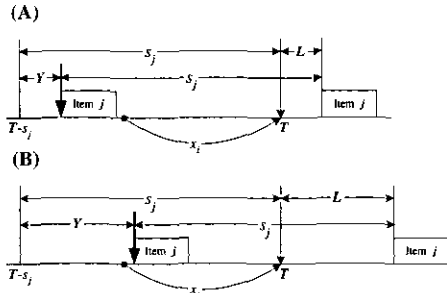


그림 7 $\delta(x_i, s_j) = 1$, (A): x_i 동안 항목 j 가 방송되지 않음 (B): x_i 동안 항목 j 가 방송됨

• Case 2. $\delta(x_i, s_j) = 1$, (그림 7 참조)

(A) Y 가 범위 $(0, s_j - x_i]$ 의 값을 가지면 항목 j 는 최종 방송 시각 조건을 만족하지 못한다.

$$\varphi(i, j, T) = 1$$

(B) Y 가 범위 $(s_j - x_i, s_j]$ 의 값을 가지면 항목 j 는 최종 방송 시각 조건을 만족하며 캐쉬의 여부는 $\Delta_m(i, j, K)$ 에 의해 결정된다.

$$\varphi(i, j, T) = \Delta_m(i, j, K)$$

(3), (4)로부터 다음을 유도할 수 있다.

$$E(\Omega_{ij}^2 | X_i = x_i \wedge \delta(x_i, s_j) = 1)$$

$$\begin{aligned} &= \int_0^{s_j - x_i} y \cdot \frac{1}{s_j} dy + \Delta_m(i, j, K) \cdot \int_{s_j - x_i}^{s_j} y \cdot \frac{1}{s_j} dy \\ &= \frac{(s_j - x_i)^2}{2s_j} + \Delta_m(i, j, K) \cdot \frac{x_i \cdot (2s_j - x_i)}{2} \end{aligned} \quad (7)$$

Case 1과 Case 2는 상호 배제이므로, $E(\Omega_{ij}^2 | X_i = x_i)$ 는 $\delta(x_i, s_j)$ 에 따라 (6) 또는 (7)이 된다. 이로부터 다음 식을 얻을 수 있다.

$$E(\Omega_{ij}^2 | X_i = x_i) = (1 - \delta(x_i, s_j)) \cdot (6) + \delta(x_i, s_j) \cdot (7) \quad (8)$$

이제 다음과 같이 Ω_{ij} 의 기대치를 구할 수 있다.

$$E(\Omega_{ij}) = E(E(\Omega_{ij} | X_i = x_i)) = \int_0^\infty E(\Omega_{ij} | X_i = x_i) \cdot P(X_i = x_i) dx_i \quad (9)$$

(8), (9)로부터

(1)과 (10)으로부터 평균 대기 시간 μ 를 얻을 수 있다. 이 분석 결과는 캐쉬 적중을 확인할 수 있는 경우만 고려하여 얻어진 것으로 실제 캐쉬의 성능은 분석 결과

$$\begin{aligned} \therefore E(\Omega_{ij}) &= (1 - \Delta_m(i, j, K)) \cdot \int_0^{s_j} \left(\frac{(s_j - x_i)^2}{2s_j} \right) \\ &\quad \cdot \lambda_i e^{-\lambda_i x_i} dx_i + \Delta_m(i, j, K) \cdot \frac{s_j}{2} \end{aligned} \quad (10)$$

보다 항상 좋다. 따라서 위 분석을 통해 얻는 μ 는 실제 평균 대기시간의 상한이 된다.

(10)을 관찰하면 다음과 같은 특징을 발견할 수 있다. 사용자 캐쉬를 가지는 방송 시스템의 평균 대기 시간은 방송 스케줄에 의해 결정되는 방송간격 s_j , 사용자 Thinking 시간 인자 λ_i , 함수 $\Delta_m(i, j, K)$ 에 의해 결정된다. 모든 (i, j) 쌍에 대하여 $\Delta_m(i, j, K) = 1$ 인 경우는 항상 캐쉬 적중에 실패하는 최악의 캐쉬라 할 수 있으며 이 때 (10)의 결과는 부록 1에서 분석한 캐쉬가 없을 때와 동일해진다. 항상 $\Delta_m(i, j, K) = 0$ 인 이상적 캐쉬가 존재하여도 평균 대기시간이 0이 되지는 않는다. 최종 방송 시각 조건 및 Thinking 시간의 제약으로 인해 불가피한 캐쉬 적중 실패가 발생하기 때문이다. (10)은 이러한 캐쉬 적중 실패에 의한 대기시간의 증가분도 반영하고 있다.

이제 대기시간의 분산 σ^2 과 관련한 식들을 구한다. 평균 대기시간 유도 방법과 비슷하게 σ^2 도 조건부 기대값을 통해 (11)과 같이 구할 수 있다. 식에서 μ 는 앞서 구한 평균 대기시간이다.

$$\sigma^2 = \left(\sum_{i=1}^M \pi_i \cdot \left(\sum_{j=1}^M b_{ij} \cdot E(\Omega_{ij}^2) \right) \right) - \mu^2 \quad (11)$$

$E(\Omega_{ij}^2)$ 도 $E(\Omega_{ij})$ 과 유사한 방법으로 구할 수 있으며 (12)와 같다. (자세한 내용은 지면 관계상 생략한다)

$$\begin{aligned} \therefore E(\Omega_{ij}^2) &= (1 - \Delta_m(i, j, K)) \cdot \int_0^{s_j} \left(\frac{(s_j - x_i)^3}{3s_j} \right) \\ &\quad \cdot \lambda_i e^{-\lambda_i x_i} dx_i + \Delta_m(i, j, K) \cdot \frac{s_j^2}{3} \end{aligned} \quad (12)$$

(1), (10), (11), (12)로부터 대기시간의 분산 σ^2 을 구할 수 있다. 분산의 경우 평균 대기시간과는 달리 최종 방송 시각 및 전체 경쟁 조건이, 분석한 값이 실제 분산 값의 상한이 되는 것을 보장해주지는 못한다. 하지만 이 분석 값은 방송 시스템에서 사용자 캐쉬를 설계할 때 이용할 수 있는 분산 값에 대한 의미 있고 유용한 척도라고 생각한다.

4. 시뮬레이션 및 결과 분석

우리는 이벤트-스케줄링 기능이 있는 간단한 시뮬레이터에서 정보 모형, 사용자 모형, 사용자 방송 스케줄 방안 등을 직접 구현하여 시뮬레이션을 수행하였다. 이 장에서는 우리가 수행한 시뮬레이션을 부분별로 설명하고, 흥미로운 방송 스케줄 및 캐쉬 정책에서 수행한 시뮬레이션 결과와 분석 결과를 비교하고 분석한다. 이를 통해 앞서 수행한 분석이 타당함을 확인하고자 한다.

4.1 실험 구성

그림 8에 실험에 사용한 데이터베이스의 정보 모형이 표시되어 있다. 모두 30개의 같은 크기(=1 단위 크기) 정보 항목으로 구성되며 기본적으로 트리 구조를 이룬다. 각 항목은 3개 이하의 자식 항목만을 가지며 전체 트리의 높이는 4이다. 개별 사용자의 첫 번째 요청 항목은 항상 "루트(항목 1)"라고 가정한다. 이어지는 해당 사용자의 요청 항목은 트리의 루트 또는 해당 사용자의 바로 전 요청 항목의 "자식" 또는 "부모" 항목으로 제한된다. 따라서 어떠한 항목도 이어 요청할 수 있는 연결 항목의 수가 5를 넘지 않는다. 그림 8에서 자식 항목으로의 전이 확률은 링크 상에 표시되며 부모 및 루트 항목으로의 전이 확률은 노드 내에 표시된다. 이 중 위쪽은 루트, 아래쪽은 부모 항목으로의 전이 확률이다. 부모와 루트가 같을 경우 하나로 표시되며, 루트는 자기 자신으로 다시 이동할 수 있다. 예를 들어 항목 5는 루트로 0.323, 부모로 0.175의 전이 확률을 가진다. $\pi = \pi^P$ 로부터 전역 요청 확률 벡터 π 를 계산할 수 있다.

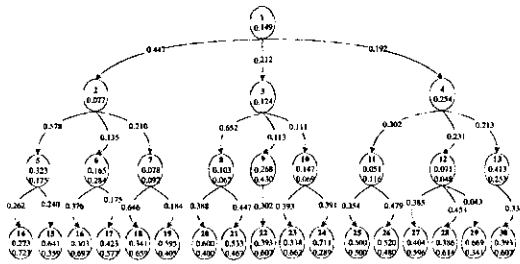


그림 8 실험에 사용한 정보 모형

단위 시간에 하나의 단위 크기를 전달할 수 있는 대역폭을 1이라고 했을 때, 실험에서 설정한 방송 채널의 대역폭은 1이다. 항목의 크기가 단위 크기라고 했으므로 하나의 항목을 방송하는데 1 단위 시간이 필요하며 중복 없이 모든 항목들을 1번 전송하는데 30 단위 시간이 필요하다.

실험에서는 그림 3의 개별 사용자 동작 모형을 가지는 10명의 사용자가 있다. 실험이 시작되면 개별 사용자들은 그림 8의 정보 모형을 따라가며 실험이 끝날 때까지 지속적으로 정보 항목을 요청한다. 충분히 긴 실험 시간 동안 이들이 생성한 요구의 대기시간을 측정하고, 각 사용자별로 대기시간의 평균과 표준편차를 구하였다.

모든 항목에 대한 Thinking 시간의 분포가 동일한 지수 분포를($\lambda_i = \lambda, 1 \leq i \leq M$) 따르도록 설정하였다.

Thinking 시간이 캐쉬의 성능에 큰 영향을 미치는 것을 고려하여 Thinking 시간의 평균 $\frac{1}{\lambda}$ 을 5부터 90 (단위 시간)까지 변화시키며 대기시간을 측정하였다. 캐쉬 크기 K는 1에서 5(단위 크기)까지 순차적으로 변화시키며 대기 시간을 측정하였다. 연결 항목의 수가 5를 넘지 않으므로, 최종 방송 시각 조건에 따라 분석에서는 캐쉬 크기가 5 보다 큰 것이 성능 향상에 도움을 주지 않는다. 하지만 실제로는 캐쉬의 크기를 5보다 크게 하는 것이 성능 향상에 도움을 줄 수 있다.

실험을 통해 살펴 본 사용자 캐싱 정책은 2가지이다. 하나는 전역 요청 확률(π_i)이 높은 것이 우선 순위를 가지는 캐싱 정책 θ 이다. 캐싱 정책 θ 는 FDM을 기반으로 하는 시스템의 대표적인 캐싱 정책이다. 다른 하나는 Thinking 중인 항목을 기준으로 전이 확률(p_{ij})이 큰 항목이 우선 순위를 가지는 캐싱 정책 ϕ 이다. 캐싱 정책 ϕ 는 LDM을 기반으로 하는 시스템의 가장 대표적인 캐싱 정책이라 할 수 있다. (이 외에도 여러 가지 캐싱 정책에 대해 실험을 하였으나 지면 관계상 그 자세한 내용을 생략한다). 분석에서 유도한 식을 이용하기 위해서는 각 정책에 대해 지시 함수 $\Delta_\theta(i, j, K)$ 와 $\Delta_\phi(i, j, K)$ 를 각각 구하여야 한다. 이들은 정보 모형 및 인자 i, j, K 를 이용하여 쉽게 구할 수 있다. 예를 들어 캐싱 정책 θ 에서 가장 큰 전역 요청 확률을 가지는 항목 1의 경우 임의의 i 및 $K \geq 1$ 에 대하여 $\Delta_\theta(i, 1, K) = 0$ 의 값을 가진다. 캐싱 정책 ϕ 의 예를 살펴 보면 $i=2, K=2$ 인 경우 j 가 5, 7인 경우만 $\Delta_\phi(i, j, K) = 0$ 이고 나머지 모든 항목에 대하여 $\Delta_\phi(i, j, K) = 1$ 이 된다.

실험한 서버 방송 스케줄 알고리즘은 2가지이다. 하나는 순환(Cyclic) 알고리즘으로 이것은 항목 1부터 M까지 순서대로 순환 방송한다. 이 알고리즘에 의한 스케줄이 등간격 속성을 가지는 것은 명백하며, 모든 항목에 대하여 방송 간격은 30(단위시간)이 된다. 순환 알고리즘은 간단하다는 장점은 있으나 항목간의 요청 확률 차이를 고려하지 않으므로 전체 성능이 좋지 않을 수 있다. 실험한 다른 알고리즘은 Vaidya 등이 제시한 알파 알고리즘이다. 이 알고리즘은 등식 (13)을 만족시키는 것을 목표로 하며, 등식 (13)이 성립할 경우 얻어지는 방송 스케줄은 등간격 속성을 가지며 항목 i 에 대하여 (14)와 같은 방송 간격을 가진다. Vaidya 등은 인자 α 값에 따라 대기 시간의 평균과 분산 사이에 타협이 이루어질 수 있음을 밝혔다. 사용자 캐쉬가 없는 상황에서 α 가 2인 경우 이 알고리즘은 평균 대기시간을 최소화

하는 스케줄을 생성한다. 그러나 α 를 3에 가깝게 할 경우 평균 대기시간이 크게 늘어나지 않고도 분산 값을 적절하게 줄일 수 있는 방송 스케줄을 얻을 수 있고, 대기시간의 분산이 중요한 성능 인자일 경우 이것이 보다 효과적이다. 우리는 α 를 2, 3 두 가지 값으로 설정하고 실험을 수행하였다.

$$s_i^q \cdot \pi_i = constant, 1 \leq i \leq M \quad (13), \quad s_i = \left(\frac{1}{\pi_i}\right)^{\frac{1}{q}} \sum_{i=1}^M \pi_i^{\frac{1}{q}} \quad (14)$$

4.2 실험 결과

이 절에서는 시뮬레이션 및 분석의 결과를 정리하고 비교한다. 결과는 크게 방송 스케줄 별로 정리하였는데, 그림 9는 순환, 그림 10은 알파-2, 그림 11은 알파-3의 결과이다. 각 그림은 4개의 그래프로 구성되는데 위로부터 각각 캐싱 정책 θ 의 대기시간 평균, 캐싱 정책 ϕ 의 대기시간 평균, 캐싱정책 θ 의 대기시간의 표준편차, 캐싱정책 ϕ 의 대기시간의 표준 편차를 나타낸다. 모든 그래프에서 Y축은 시간 값이며, X축은 Thinking 시간 인자에 대한 것으로 왼쪽부터 $\frac{1}{\lambda}$ 이 5, 10, 20, 30, 40, 50, 90 이다. 분석 값과 시뮬레이션 결과 값은 점선과 실선으로 구분되며, 캐쉬 크기에 따른 구분은 선구분자(그림의 범례 참조)로 한다.

우리는 분석을 통해 얻은 평균 대기시간이 실제 평균 대기시간의 상한이라고 주장하였다. 그리고 분석을 통해 얻은 대기시간의 분산은 실제 대기시간 분산의 상한임을 보장하지는 못하지만 의미 있고 유용한 척도라고 주장하였다. 그림 9, 10, 11에서 실선과 점선의 관계를 살펴보면 이러한 주장이 타당함을 알 수 있다. Thinking 시간이 길수록 분석 결과와 시뮬레이션 결과의 차이가 줄어드는 현상을 확인할 수 있는데 이것은 최종 방송 시작 조건이 상당 수준 완화되기 때문이라고 파악된다.

4.3 토론

그림 9, 10, 11의 실험결과에서 다음과 같은 몇 가지 특징을 관찰할 수 있었다.

- LDM 기반 캐싱 정책 ϕ 가 FDM 기반 캐싱 정책 θ 에 비해 훨씬 좋은 성능을 가진다.
- 일반적으로 캐싱 정책 ϕ 에서는 사용자 Thinking 시간이 길수록 성능이 향상되었지만, 캐싱 정책 θ 에서는 사용자 Thinking 시간의 변화가 성능에 영향을 거의 주지 않는다.

2) 실제 논문[2]에서 제시한 등식 $\frac{s_i^q \cdot \pi_i}{l_i} = constant, 1 \leq i \leq M$ 이다. 여기서 l_i 는 항목의 크기를 나타낸다. 우리는 가정에서 모든 항목이 같은 크기라고 하였으므로 위 등식에서 l_i 를 생략하였다.

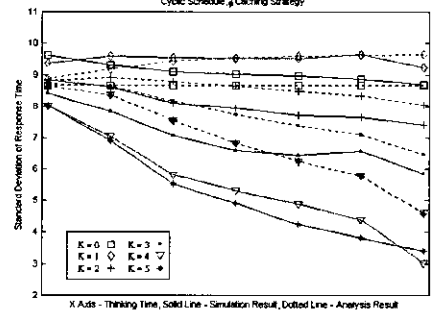
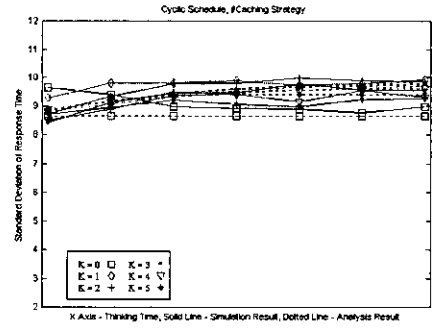
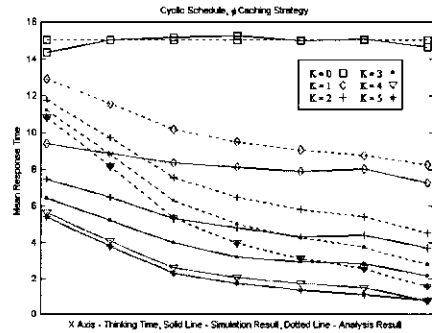
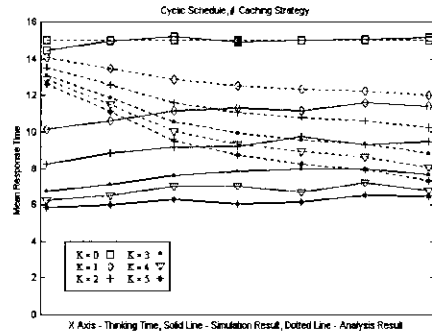


그림 9 Cyclic 스케줄에 대한 캐싱 정책별 대기시간의 평균 및 표준편차

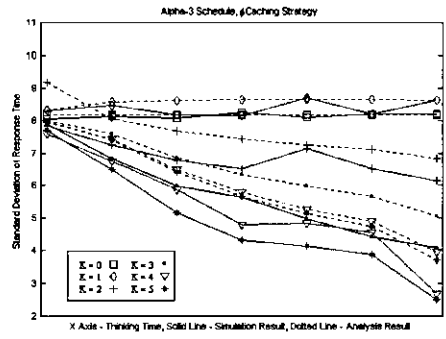
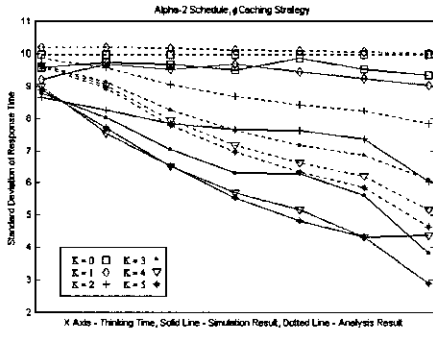
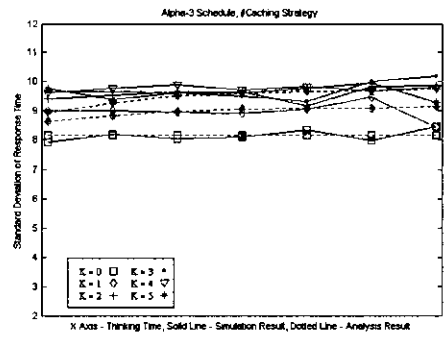
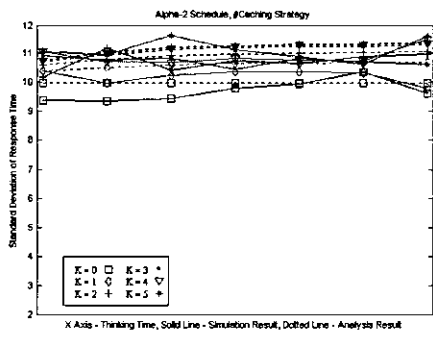
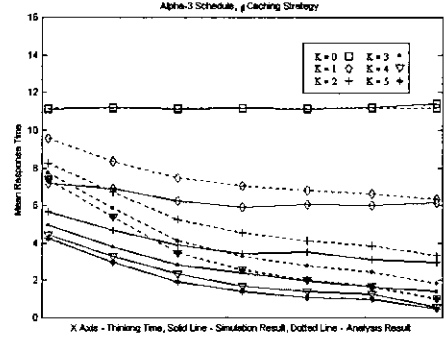
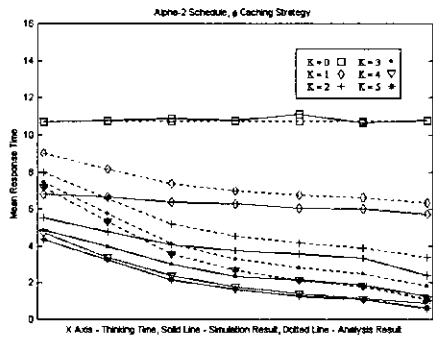
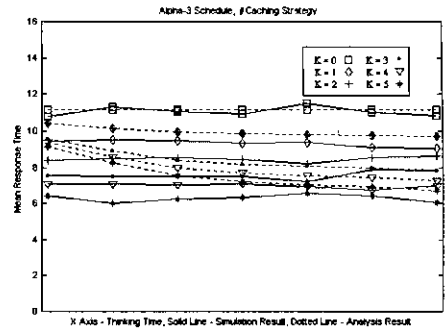
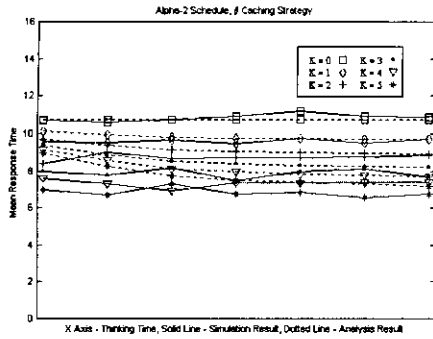


그림 10 $\alpha-2$ 스케줄에 대한 캐싱 정책별 대기시간의 평균 및 표준편차

그림 11 $\alpha-3$ 스케줄에 대한 캐싱 정책별 대기시간의 평균 및 표준편차

- 캐싱 정책 ϕ 에서 캐쉬 크기를 1에서 2, 2에서 3으로 변화시킬 때 얻는 성능 향상에 비해 3 이상의 크기로 증가시켜서 얻는 성능 향상의 미미하였다.
- 캐싱 정책 ϕ 의 경우 사용자 캐쉬는 평균 대기시간 뿐 아니라 대기시간의 분산도 줄이는 역할을 한다. 그러나 캐싱 정책 θ 의 경우 대기 시간의 분산이 더욱 나빠졌다.
- 캐싱 환경이 좋을 경우, 즉 캐쉬의 크기가 크고 사용자 Thinking 시간이 길 경우 캐싱 정책 ϕ 는 평균 대기 시간이 0에 근접할 정도의 우수한 성능을 보여준다. 시뮬레이션 결과로만 평가하면 캐싱 환경이 나쁜 경우에도 캐싱 정책 ϕ 에 의한 성능 향상, 엄밀하게는 평균 대기시간의 감소는 상당히 크게 나타났다. (실험 결과 $K=1$, $\frac{1}{\lambda}=5$ 인 경우에도 평균 대기시간은 캐쉬가 없는 경우에 비해 30% 이상 줄어들었다.) 이는 캐싱 정책 ϕ 가 LDM의 특징을 잘 이용하여 사용자의 미래 요구를 효과적으로 예측하였기 때문이다.
- 순환 알고리즘에 비해 알파 알고리즘은 캐쉬가 있을 때도 대기시간의 평균뿐만 아니라 분산 측면에서도 대체로 우수하다.
- 알파 알고리즘에서 $\alpha=3$ 으로 설정하는 것이 캐쉬의 유, 무에 관계없이 대기시간의 분산 측면에서는 $\alpha=2$ 인 경우보다 우수하다.
- 알파 알고리즘에서 캐쉬가 없을 경우 $\alpha=2$ 로 설정하면 평균 대기시간을 최소화할 수 있음이 알려져 있다. 그런데 캐쉬가 있고 캐싱 환경이 좋을 경우, $\alpha=3$ 이 평균 대기시간 측면에서도 보다 나은 선택임을 관찰할 수 있다. 캐쉬가 있어도 캐싱 환경이 좋지 않을 경우는 $\alpha=2$ 인 경우가 여전히 유리하다.
- 최적의 캐쉬 크기를 찾는 것은 아주 어려운 일이지만, 그것이 연결 항목의 개수와 밀접한 연관 관계가 있음을 알 수 있다.

5. 결론

우리는 이 논문을 통해 LDM로 정의한, 데이터 방송 시스템의 정보 데이터베이스를 위한 새로운 정보 모형을 제시하였다. 이 정보 모형의 특징을 잘 이용하면 사용자 요구의 효과적 예측이 가능해져 사용자 캐쉬를 통해 큰 성능 향상을 얻을 수 있다. 이러한 성능 향상을 정량화하는 것은 사용자 캐쉬의 크기와 캐싱 정책의 설계 및 구현에 꼭 필요한 것이라고 할 수 있다. 그런데 방송 시스템의 성능은 캐쉬뿐 아니라 방송 스케줄에 의

해서도 큰 영향을 받는다. 따라서 방송 스케줄, 사용자 캐쉬 크기, 캐싱 정책 등 방송 시스템의 여러 성능 요인을 함께 고려하는 이론적 성능 평가 도구가 필요하다.

우리는 이 논문에서 LDM과 사용자 캐쉬를 가지는 데이터 방송 시스템에서 사용자 대기시간의 평균 및 분산을 분석하였다. 캐쉬 적중에 대해 최악 분석 기법을 적용하여 대기시간의 평균 및 분산과 관련한 두 개의 주요한 식을 유도할 수 있었다. 식의 하나는 평균 대기시간의 상한을, 다른 하나는 엄밀하게 보장할 수는 없지만 대기시간 분산의 유용한 근사치를 구할 수 있도록 해준다. 이 식들은 특정 방송 방안이나 캐싱 방안에 종속된 것이 아니다. 따라서 방송 시스템의 여러 구성 방안들간의 성능 비교를 위한 이론적 도구로서의 역할을 수행할 수 있다. 우리는 분석의 타당성 검토를 위해 다양한 설정 아래, 심도 있는 시뮬레이션을 수행하였다. 분석 결과와 시뮬레이션 결과의 비교, 검토를 통해 수행한 분석의 타당성을 입증하였다.

관련하여 앞으로 수행할 연구 방향은 다음과 같다. 시스템을 구성하는 각 부분에 대해, 보다 일반적인 가정 및 다양한 상황 아래 대기 시간 분석이 필요하다. 예를 들면 항목의 크기가 다른 경우, 사용자마다 항목간의 전이 확률이 다를 경우, 여러 개의 방송 채널이 있을 경우, 사용자 요청 취소를 고려한 사용자 모형을 채택할 경우 등을 들 수 있다. 나아가 사용자 캐쉬를 고려할 경우 사용자 대기 시간을 최적화 할 수 있는 방안에 관한 연구가 필요하다.

참고 문헌

- [1] N. H. Vaidya and S. Hameed, "Scheduling Data Broadcast In Asymmetric Communication Environments," to appear in ACM/Baltzer Wireless Networks (WINET) journal
- [2] Shu Jiang and Nitin Vaidya, "Response Time in Data Broadcast Systems: Mean, Variance and Trade-Off," International Workshop on Satellite-based Information Services (WOSBIS), 1998
- [3] Mostafa H. Ammar, "Response Time in a Teletext System : An Individual User's Perspective," Vol 35, No. 11. November 1987
- [4] Chi-Jiun Su and Leandros Tassioulas, "Joint Broadcast Scheduling and User's Cache Management for Efficient Information Delivery," ACM Journal on Wireless Networks, 1999.
- [5] Leandros Tassioulas and Chi-Jiun Su, "Optimal Memory Management Strategies For Mobile Computing," IEEE JSAC Special Issue on Networking and Performance Issues of Personal

Mobile Communications, Vol. 15, No. 7, Page 1224-1238, September, 1997.

- [6] Swarup Acharya, Michael. Franklin, and Stanley. Zdonik, "Prefetching from a Broadcast Disk," International Conference on Data Engineering, Feb 1996.
- [7] Demet Aksoy, Michael Franklin, "Scheduling for Large-Scale On-Demand Data Broadcasting," IEEE INFOCOM 1998.

부록 1. 캐쉬가 없을 경우의 대기시간의 평균 및 분산

캐쉬가 없을 경우 함수 $\rho(i, j, T)$ 는 항상 1(=캐쉬 적중 실패)의 값을 가지므로 (4)에서 $\Omega_{ij} = Y$ 가 된다. (3)을 적용하여 $E(Y) = \frac{s_i}{2}$ 를 쉽게 구할 수 있으며, 평균 대기시간 μ 를 다음과 같이 유도할 수 있다.

$$\mu = \sum_{i=1}^M \pi_i \cdot \left(\sum_{j=1}^M \rho_{ij} \cdot E(\Omega_{ij}) \right) = \sum_{i=1}^M \pi_i \cdot \left(\sum_{j=1}^M \rho_{ij} \cdot \frac{s_j}{2} \right) = \sum_{i=1}^M \frac{s_i}{2} \cdot \left(\sum_{j=1}^M \pi_j \cdot \rho_{ij} \right)$$

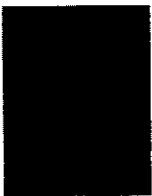
LDM의 이산 마코프 연쇄 가정에 의해 $\pi = \pi P$ 이므로 $\pi_j = \sum_{i=1}^M \pi_i \cdot \rho_{ij}, (1 \leq j \leq M)$ 가 성립한다. 따라서

$$\mu = \sum_{i=1}^M \frac{s_i}{2} \cdot \left(\sum_{j=1}^M \pi_j \cdot \rho_{ij} \right) = \sum_{i=1}^M \pi_i \cdot \frac{s_i}{2} \tag{A-1}$$

Vaidya 등은 등간격 속성 가정 아래 사용자 캐쉬가 없을 경우의 대기시간 평균 $\hat{\mu}$ ((A-2))와 분산 $\hat{\sigma}^2$ ((A-3))을 구하였다. (A-1)과 (A-2)가 동일함은 쉽게 알 수 있다. 대기시간의 분산에 대해서도 같은 결과를 얻을 수 있다. 이를 통해 우리의 분석이 기존 연구 결과에 배치되지 않음을 알 수 있다.

$$\hat{\mu} = \frac{1}{2} \sum_{i=1}^M s_i \pi_i \tag{A-2}$$

$$\hat{\sigma}^2 = \frac{1}{3} \sum_{i=1}^M \pi_i s_i^2 - \hat{\mu}^2 \tag{A-3}$$



김 중 덕

1994년 서울대학교 전산과학과 졸업(학사). 1996년 서울대학교 전산과학과 졸업(석사). 1996년 ~ 현재 서울대학교 컴퓨터공학부 박사과정.



김 중 권

1981년 서울대학교 산업공학과 졸업(학사). 1987년 일리노이 주립대학 졸업(박사). 1987년 ~ 1991년 벨코어 연구소 연구원. 1991년 ~ 현재 서울대학교 전산과학과 부교수 재직