

상호참조시간을 고려한 단기간 임시지역성 측정

김 영 일[†]·심 재 홍^{††}·최 경 희^{†††}·정 기 현^{††††}

요 약

임시지역성은 웹 캐싱전략에서 고려해야 할 중요한 요소 중의 하나이며, 이를 정확히 측정하는 것은 매우 중요하다. 임시지역성을 측정하기 위한 다양한 방법이 제시되었으나, 상호 참조시간을 고려하지 않는 스택을 이용한 측정도구를 사용함으로써 정확한 임시지역성 측정에 한계가 있었다. 본 연구에서는 상호참조시간을 고려한 단기간 임시지역성 측정도구를 제안하고, NLANR과 NASA의 로그 파일로 실험한 결과를 논의한다. 실험 결과 스택기반의 기존 측정도구보다 상호참조시간 기반의 제안 측정도구가 단기간 임시지역성을 보다 정확히 측정할 수 있었다.

Measurement of Short-term Temporal Locality Based on Request Interarrival Time

Yeong-Il Kim[†] · Jae-Hong Shim^{††} · Kyung-Hee Choi^{†††} · Gi-Hyun Jung^{††††}

ABSTRACT

Temporal locality of Web server references is one of the important characteristics to be considered in the design of a Web caching strategy, and it is important to measure the temporal locality exactly. Various methods to estimate the temporal locality have been proposed, however, Web server designers have still troubles in its measurement by using the tools that don't reflect the interarrival time of document requests. In this paper, we propose a measurement tool for short-term temporal locality based on request interarrival time, and discuss the simulation results based on the traces from NLANR and NASA Web sites. The results show that the proposed tool estimates the short-term temporal locality more exactly than that based on a stack.

키워드: 웹 캐싱(Web Caching), 웹 참조(Web Reference), 임시지역성 측정(Temporal Locality Measurement), 단기간 임시지역성(Short-Term Temporal Locality)

1. 서 론

임시지역성(temporal locality)이란 웹 사용자에게 의해 한번 참조된 문서가 가까운 미래에 다시 참조되는 것을 말한다. 임시지역성은 문서의 인기도에 의해 발생하는 장기간 임시지역성과 문서 참조의 상호 관계에 의해 발생하는 단기간 임시지역성이 있다. 이러한 임시지역성은 웹 워크로드의 특성화, 복제, 선인출 기법(prefetching protocol) 등은 물론, 웹 성능의 향상을 위한 캐싱전략에 커다란 영향을 미치고 있다[1, 2, 6-10]. 웹 사용이 증가되면서 웹 접근에 대한 참조시간의 지연을 줄이고 네트워크 대역폭을 절약하려는 연구들이 진행되었고, 이러한 연구 중의 하나가 캐싱이다[3, 13-15]. 또한 웹 접근 패턴의 특성화에서 참조의 임시지역성이 매우 강하게 존재한다는 연구 결과가 제시되었다[1]. 따라서 임시지역성은 캐싱전략에서 고려되어야 할 중

요한 요소 중의 하나로 이를 정확히 측정하는 것은 매우 중요하다.

기존 연구에서 임시지역성을 측정하기 위한 다양한 방법이 제안되었다[1-4]. Cao[4]는 상호 도착시간을 활용한 RIM(Reference Interarrival Model)[11]을 사용하여 임시지역성을 측정하였다. 이 방법은 문서 요청흐름에서 장기간 임시지역성을 측정하기에는 유용하다. 그러나 인기도와 문서참조의 상호관계를 분리하지 않기 때문에, 인기도가 높을 경우 상호관계가 낮음에도 불구하고 상호관계가 높은 것으로 측정될 수 있는 단점이 있다. Jin[2]은 한번 참조된 문서가 다시 참조되는 비율을 로그-로그 그래프(X축: 재참조 간격, Y축: 재참조 비율)로 나타내었다. 이 그래프에서 Zipf-law [5] 파라미터 α 를 산출하였고, 인기도를 균등하게 하여 Zipf-law 파라미터 β 를 산출하였다. 이 파라미터들을 이용하여 장기간 임시지역성과 단기간 임시지역성을 분리하여 측정하였다. 이 방법은 요청된 문서 전체에 대한 단기간 임시지역성을 측정하고자 할 경우 유용하다. 그러나 문서 하나 하나의 단기간 임시지역성을 측정하기는 어렵다.

Almeida[1]는 LRUSM(Least Recently Used Stack Model)

[†] 준 회원: 아주대학교 대학원 컴퓨터공학과

^{††} 정 회원: 조선대학교 인터넷소프트웨어공학부 교수

^{†††} 정 회원: 아주대학교 정보통신전문대학원 교수

^{††††} 정 회원: 아주대학교 전자공학부 교수

논문접수: 2003년 5월 19일, 심사완료: 2003년 12월 31일

[12]을 사용하여 문서들의 임시지역성을 측정하였다. LRUSM에서는 임시지역성을 문서 요청흐름의 평균 스택깊이로 판단한다. 참조된 문서의 스택깊이는 참조될 당시 그 문서가 있었던 스택의 길이로 정의된다. 한번 참조된 문서는 스택의 상단으로 옮겨지고 참조된 문서의 스택깊이보다 낮은 문서들은 스택깊이가 1씩 증가한다. 요청된 전체 문서들의 평균 스택깊이는 각 문서들의 요청 당시 스트림의 스택깊이의 평균값이다. 원시 스트림과 그들의 요청 순서를 임의로 바꾼 스트림에 대해 평균 스택깊이를 측정하였을 때, 요청 순서를 임의로 바꾼 스트림보다 원시 스트림의 평균 스택깊이가 낮으면 임시지역성이 높고, 평균 스택깊이가 깊으면 임시지역성이 낮은 것으로 판단한다. 이 방법은 LRU (Least Recently Used)방식의 캐싱전략을 사용하는 웹서버에 유리하다. 그러나 인기도와 단기간의 임시지역성을 분리하는데 어려움이 있어서 인기도에 의해 나타난 값은 단기간의 임시지역성으로 오인할 수 있는 단점이 있다.

Mahanti[3]는 인기도와 단기간 임시지역성을 분리하여 측정하였다. 인기도에 의한 장기간 임시지역성은 인기도가 높은 문서들의 집합(hot set)이 1일 간격으로 겹치는 비율을 계산하여 간단히 측정하였다. 단기간 임시지역성 측정을 위해선 LRUSM이 인기도가 높은 문서들의 집합을 분리할 수 없는 단점을 보완하여 새로운 측정도구 T_{ij} 를 제안하였다. T_{ij} 는 LRU 캐싱전략에서 문서 D_i 가 스택 j 에 있을 때 전체 참조된 수의 비율에 대해 문서 D_j 가 스택 j 에 있을 때 문서 D_i 가 참조된 수의 비율과의 상대적인 비율을 말한다. 이 측정법은 특정 스택깊이에서 재 참조되는 문서들을 분석하여 측정값이 1보다 크면 클수록 단기간 임시지역성이 높은 것으로 판단하고, 1보다 작으면 작을수록 단기간 임시지역성이 낮은 것으로 판단한다. 이는 각 문서별로 단기간 임시지역성을 측정할 수 있는 이점이 있다. 그러나 시간적으로 잠시 후에 참조될 문서가 많은 다른 문서들의 참조로 인해 스택깊이가 깊어질 경우 단기간 임시지역성이 낮은 것으로 측정되거나, 또는 워크로드가 적은 경우 단기간 임시지역성이 낮은데도 불구하고 높은 것으로 측정될 수 있는 단점이 있다.

따라서 본 연구에서는 Mahanti[3]의 단기간 임시지역성 측정도구 T_{ij} 의 단점을 보완한 새로운 측정도구 M_{iu} 를 제안하고자 한다. M_{iu} 는 요청흐름의 임시지역성에서 인기도에 의한 영향을 분리하여 측정할 수 있으며, 시간적으로 잠시 후에 참조될 문서가 다른 문서의 많은 참조로 인하여 단기간 임시지역성이 낮은 것으로 오인되는 것을 줄일 수 있는 단기간 임시지역성 측정도구이다.

본 논문의 구성은 다음과 같다. 2장에서는 스택기반의 단기간 임시지역성 측정도구에 관한 기존 연구를 분석한다. 3장에서는 단기간 임시지역성 특성을 보다 명확하게 측정할 수 있는 상호참조시간을 고려한 단기간 임시지역성 측정도구를 제안한다. 4장에서는 기존 웹 사이트 추적자료를 이용

한 다양한 실험결과를 제시하고, 이의 분석을 통해 제안된 측정도구의 타당성을 검증한다. 마지막 5장에서 결론 및 향후 연구로 마무리한다.

2. 관련 연구 및 배경

단기간 임시지역성을 측정하는 측정도구 T_{ij} 는 임의의 문서 D_i 가 스택깊이 j 에 있을 때 전체 참조 수의 비율에 대한 D_j 의 참조 비율로 정의된다[3].

$$T_{ij} = \frac{\text{문서 } D_i \text{가 스택 } j \text{에 있을 때 문서 } D_i \text{의 참조비율}}{\text{문서 } D_i \text{가 스택 } j \text{에 있을 때 전체 문서들의 참조비율}}$$

즉,

$$T_{ij} = \frac{a_{ij} / R_i}{c_{ij} / T} \text{이다.}$$

여기서, $T = \sum R_i$: 전체 문서 요청 횟수

R_i = 문서 D_i 의 요청 횟수

a_{ij} = 문서 D_i 가 스택 j 에 있을 때, 문서 D_j 의 요청 횟수

c_{ij} = 문서 D_i 가 스택 j 에 있을 때, 총 요청 횟수이다.

예를 들어, 문서 D_1, D_2, D_3 가 $D_1 D_1 D_2 D_2 D_1 D_1 D_3 D_2 D_3 D_3 D_1 D_1 D_3 D_2 D_3$ 의 순서로 요청되었고, 캐싱전략으로 LRU를 사용할 때, $R_1 = 6, R_2 = 4, R_3 = 5$ 이며, $T = 15$ 이다. 이때 $a_{ij}(i = 1, 2, 3, j = 1, 2, \dots)$ 및 $c_{ij}(i = 1, 2, 3, j = 1, 2, \dots)$ 의 변화되는 값은 <표 1>과 같다.

<표 1>의 인자 값으로 T_{ij} 를 구하면,

$$T_{11} = \frac{3/6}{5/15} = 1.33, \quad T_{12} = \frac{1/6}{4/15} = 0.63,$$

$$T_{13} = \frac{1/6}{4/15} = 0.63, \quad T_{21} = \frac{1/4}{4/15} = 0.94$$

$$T_{22} = \frac{0/4}{4/15} = 0.00, \quad T_{23} = \frac{2/4}{4/15} = 1.88$$

$$T_{31} = \frac{1/5}{4/15} = 0.75, \quad T_{32} = \frac{3/5}{4/15} = 2.25,$$

$$T_{33} = \frac{0/5}{0/15} = NA \text{이다.}$$

이 측정값들은 임의의 스택깊이에서 특정 문서의 단기간 임시지역성이 '높다, 낮다'를 판단하는 기준이 된다. T_{ij} 값이 1보다 크면 단기간 임시지역성이 높다고 판단하고, 1보다 작으면 단기간 임시지역성이 낮다고 판단한다. 위의 문서 중 D_1 의 T_{11} 은 1.33으로 단기간 임시지역성이 높고, D_3 의 T_{31} 은 0.75로 단기간 임시지역성이 낮다고 판단한다.

Mahanti[3]는 USask(Saskatchewan 대학의 서버)의 로그 데이터와 NLANR(National Laboratory for Applied Network Research)의 로그 데이터 중 각각 인기도가 높은 25개의 문서에 대하여 T_{ij} 를 측정하였다. USask의 로그 데이터 중 일

〈표 1〉 측정도구 T_{ij} 의 인자 값

순서		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	최종값
요청문서		D_1	D_1	D_2	D_2	D_1	D_1	D_3	D_2	D_3	D_3	D_1	D_1	D_3	D_2	D_3	
스택	1	D_1	D_1	D_2	D_2	D_1	D_1	D_3	D_2	D_3	D_3	D_1	D_1	D_3	D_2	D_3	
	2			D_1	D_1	D_2	D_2	D_1	D_3	D_2	D_2	D_3	D_3	D_1	D_3	D_2	
	3							D_2	D_1	D_1	D_1	D_2	D_2	D_2	D_1	D_1	
D_1 의 c_{ij}, a_{ij}	c_{11}		1				2	3					4	5			5
	c_{12}				1	2			3						4		4
	c_{13}									1	2	3				4	4
	a_{11}		1				2						3				3
	a_{12}					1											1
	a_{13}											1					1
D_2 의 c_{ij}, a_{ij}	c_{21}				1	2				3						4	4
	c_{22}						1	2			3	4					4
	c_{23}								1				2	3	4		4
	a_{21}				1												1
	a_{22}																0
	a_{23}								1						2		2
D_3 의 c_{ij}, a_{ij}	c_{31}								1		2	3			4		4
	c_{32}									1			2	3		4	4
	c_{33}																0
	a_{31}										1						1
	a_{32}									1				2		3	3
	a_{33}																0

부 문서들은 낮은 스택깊이에서는 T_{ij} 가 높게 나타나다가, 스택의 깊이가 깊어짐에 따라 T_{ij} 가 작아지는 것이 관측되었다. 이는 문서의 요청수가 많아짐에 따라 재 참조되는 수가 상대적으로 적어진다는 것을 의미한다. 따라서 낮은 스택깊이에서는 단기간 임시지역성이 높다고 판단하였다. NLANR 로그 데이터의 실험에서도 이와 비슷한 T_{ij} 값들이 측정되었다.

이러한 측정방법은 스택을 기반으로 하였기 때문에 워크로드의 많고 적음에 따라 단기간 임시지역성의 높고 낮음을 오인할 수 있다. 즉, 시간적으로 잠시 후에 다시 참조될 문서들이 많은 다른 문서들의 참조로 인하여 평균 스택깊이가 깊어져 단기간 임시지역성이 낮은 것으로 오인될 수 있다. 반면에 상당한 시간이 지난 후에 참조될 문서들이 워크로드가 적은 경우 평균 스택깊이가 낮아져 단기간 임시지역성이 높은 것으로 오인될 수 있다. 예를 들어, 문서 D_1 의 T_{11} 은 1.33으로 1보다 커서 단기간 임시지역성이 높다고 판단했지만, 워크로드가 적어서 시간상으로 멀리 떨어져 있다면 단기간 임시지역성이 낮아 질 수 있는 경우이다. 반면 문서 D_3 의 T_{31} 은 0.75로 1보다 작아 단기간 임시지역성이 낮은 것으로 판단했지만, 워크로드가 많아 시간상으로 가까이 있다면 단기간 임시지역성이 높아 질 수 있는 경우이다. 따라서 스택을 기반으로 하는 단기간 임시지역성 측정방법은 시간개념을 고려하지 않기 때문에 워크로드의 양에 따라 측정값이 달라질 수 있으며, 이로 인해 정확한 단기간 임시지역성을 측정하는 데는 한계가 있다.

3. 상호참조시간을 고려한 임시지역성 측정도구

스택을 이용한 측정방법은 임시지역성의 정의에서 나타나는 시간개념이 고려되지 않았다. 따라서 시간적으로 잠시 후에 다시 참조될 문서를 중간에 다른 문서가 많이 참조된다고 해서 캐시에서 제거한다면, 이는 곧 문서 참조 성공률을 감소시키는 한 요인이 될 수 있다. 또한 문서 요청흐름에서 참조된 상호 시간적 관계를 그래프로 관찰하는 방법은 각각의 문서에 대한 임시지역성을 측정하는데 어려움이 있다. 본 연구에서는 시간개념을 도입하여 각각의 문서가 어느 시간대에 단기간 임시지역성의 특성이 있는지 명확하게 측정할 수 있는 도구 M_{iu} 를 제안한다. 또한 이를 이용하여 문서 요청흐름에서의 단기간 임시지역성을 측정하고자 한다.

측정도구 M_{iu} 는 참조된 문서 총수에 대한 단위시간 내에 요청된 문서 수의 비율과 어떤 문서 D_i 가 참조된 횟수에 대한 그 문서의 단위시간 내에 요청된 수의 비율과의 상대적인 비율로 정의된다. 단위시간은 단기간 임시지역성을 측정하기 위한 적정한 시간으로 단위시간의 변화에 따라 M_{iu} 의 값에 영향이 미칠 것으로 예상된다.

$$M_{iu} = \frac{\text{문서 } D_i \text{가 참조된 횟수에 대한 문서 } D_i \text{가 } u \text{ 단위시간 내에 재참조된 횟수의 비율}}{\text{참조된 문서 총횟수에 대한 } u \text{ 단위시간 내에 참조된 횟수의 비율}}$$

즉,

$$M_{iu} = \frac{b_{iu} / R_i}{d_{iu} / T} \text{ 이다.}$$

여기서, T = 총 문서 요청 횟수
 R_i = 문서 D_i 의 요청 횟수
 b_{iu} = 문서 D_i 가 요청된 후, u 단위시간 내에 재 요청된 횟수
 d_{iu} = 문서 D_i 가 요청된 후, u 단위시간 내에 재 요청되었을 때 그 사이에 D_i 를 포함한 요청 횟수이다.

$$M_{13} = \frac{2/6}{6/15} = 0.83, \quad M_{21} = \frac{1/4}{1/15} = 3.75$$

$$M_{22} = \frac{0/4}{0/15} = 0.00, \quad M_{23} = \frac{0/4}{0/15} = 0.00$$

$$M_{24} = \frac{1/4}{4/15} = 0.94, \quad M_{25} = \frac{1/4}{6/15} = 0.63$$

$$M_{31} = \frac{2/5}{3/15} = 2.00, \quad M_{32} = \frac{1/5}{2/15} = 1.50,$$

$$M_{33} = \frac{1/5}{3/15} = 1.00 \text{ 이다.}$$

2장의 예제에서는 문서 요청흐름에서 참조된 순서만을 고려하여 T_{ij} 를 측정하였다. 여기서는 요청흐름에서 순서는 물론 시간도 함께 고려하여 측정하고자 한다. <표 2>는 2장에서 예를 든 요청흐름의 순서와 같으나, 요청된 시간이 포함된 단위시간별 요청흐름이다. 요청흐름에서 <표 2>와 같은 요청시각으로 참조되었을 때, $R_1 = 6, R_2 = 4, R_3 = 5$ 이고, $T = 15$ 이다. 이때 $b_{iu}(i = 1, 2, 3, u = 1, 2, \dots)$ 및 $d_{iu}(i = 1, 2, 3, u = 1, 2, \dots)$ 의 변화되는 값은 <표 2>와 같다. D_1 의 요청간격에서 $1u$ 는 없으므로 $b_{11} = 0, 2u$ 는 3개이므로 $b_{12} = 3, 3u$ 는 2개이므로 $b_{13} = 2$, 문서 D_1 이 요청되고 $1u$ 내에 다시 D_1 이 요청될 때까지 요청된 것은 없으므로 $d_{11} = 0, 2u$ 내에 다시 요청된 수는 3과 2이므로 $d_{12} = 5, 3u$ 내에 다시 요청된 수는 1과 5이므로 $d_{13} = 6$ 이다. 같은 방법으로 문서 D_2, D_3 에 대한 b_{iu} 및 d_{iu} 값을 구할 수 있다.

<표 2>의 인자 값으로 측정도구 M_{iu} 의 값을 구하면,

$$M_{11} = \frac{0/6}{0/15} = 0.00, \quad M_{12} = \frac{3/6}{5/15} = 1.50,$$

앞서 측정된 T_{ij} 와 M_{iu} 를 <표 3>과 같이 비교해 보면 차이가 있음을 알 수 있다. 문서 D_1 의 T_{11} 은 1.33이고, M_{11} 은 0.00이다. 이는 재 참조되는 사이에 워크로드가 적어서 낮은 스택 값을 갖게 되고, 이로 인하여 단기간 임시지역성이 높은 것처럼 보이지만 시간상으로 보면 상당한 시간이 지난 후에 재 참조되는 것을 알 수 있다. 이는 T_{ij} 로 단기간 임시지역성을 측정할 때, 워크로드가 적을 경우 단기간 임시지역성이 없는데도 있는 것처럼 오인할 수 있다는 것을 보여준다. 또한 문서 D_3 의 T_{31} 은 0.75이고, M_{31} 은 2.00이다. 이는 재 참조되는 사이에 다른 문서들이 참조되어 높은 스택 값을 갖게 되고 이로 인하여 단기간 임시지역성이 낮은 것처럼 보이지만, 시간상으로 보면 가까운 시간 내에 재 참조되는 것을 알 수 있다. 이는 T_{ij} 로 단기간 임시지역성을 측정할 때 워크로드의 상태에 따라 단기간 임시지역성이 있는데도 없는 것으로 오인할 수 있다는 것을 보여준다. 그러나 측정도구 M_{iu} 는 요청 순서는 물론 요청 시간까지 고려하여 시간적으로 가까운 시간 내에 다시 요청되는 것을 정확히 측정할 수 있다. 다만, u 값을 충분한 시간간격으로

<표 2> 단위시간별 요청흐름과 측정도구 M_{iu} 의 인자 값($u = 1$ 초)

연 번	문 서	요청시각 (초)	D_1 의 요청간격	d_{1u}	D_2 의 요청간격	d_{2u}	D_3 의 요청간격	d_{3u}
1	D_1	1	첫요청					
2	D_1	3.5	$3u$	$d_{11} = 1$				
3	D_2	4			첫요청			
4	D_2	5			$1u$	$d_{21} = 1$		
5	D_1	5.5	$2u$	$d_{12} = 3$				
6	D_1	7.5	$2u$	$d_{12} = 4$				
7	D_3	8.5					첫요청	
8	D_2	9			$4u$	$d_{24} = 4$		
9	D_3	9.5					$1u$	$d_{31} = 2$
10	D_3	10					$1u$	$d_{31} = 3$
11	D_1	10.5	$3u$	$d_{13} = 6$				
12	D_1	12	$2u$	$d_{13} = 5$				
13	D_3	13					$3u$	$d_{33} = 3$
14	D_2	14			$5u$	$d_{25} = 6$		
15	D_3	15					$2u$	$d_{32} = 2$
최종값	15		$b_{11} = 0$ $b_{12} = 3$ $b_{13} = 2$	$d_{11} = 0$ $d_{12} = 5$ $d_{13} = 6$	$b_{21} = 1$ $b_{24} = 1$ $b_{25} = 1$	$d_{21} = 1$ $d_{24} = 4$ $d_{25} = 6$	$b_{31} = 2$ $b_{32} = 1$ $b_{33} = 1$	$d_{31} = 3$ $d_{32} = 2$ $d_{33} = 3$

하게 되면 M_{iu} 값이 계산 불가능하거나 0이 되는데 이 시점에서는 단기간 임시지역성이 없다는 것을 의미한다.

〈표 3〉 T_{ij} 와 M_{iu} 의 측정 값 비교

문서	1bucket 또는 1u		2bucket 또는 2u		3bucket 또는 3u	
	T_{ij}	M_{iu}	T_{ij}	M_{iu}	T_{ij}	M_{iu}
D_1	1.33	0.00	0.63	1.50	0.63	0.83
D_2	0.94	3.75	0.00	0.94	1.88	0.63
D_3	0.75	2.00	2.25	1.50	0.00	1.00

4. 실험 및 평가

단기간 임시지역성 측정도구인 M_{iu} 의 타당성을 검증하기 위하여 NLANR과 NASA의 로그 파일을 실험 데이터로 사용하여 M_{iu} 의 성능을 측정하였다. 먼저, 워크로드와 M_{iu} 값과의 상관관계를 알아보기 위해 단위시간을 변경하여 워크로드에 변화를 주어 실험하고 결과를 분석한다. 임시지역성은 인기도에 의한 것과 문서참조의 상호관계에 의한 것으로 나누어진다. 이를 구별하기 위해 두 번째 실험에서는 인도가 높은 문서들과 낮은 문서들의 M_{iu} 값을 측정하여 인기도와 M_{iu} 값의 상관관계를 분석한다. 세 번째 실험에서는 요청 흐름에 따른 문서들의 M_{iu} 측정값과 참조순서를 무작위순서로 변경한 후 측정된 M_{iu} 값을 비교함으로써 M_{iu} 가 단기간 임시지역성을 적절히 반영하는지를 확인하고자 한다.

4.1 워크로드와 M_{iu} 값과의 관계

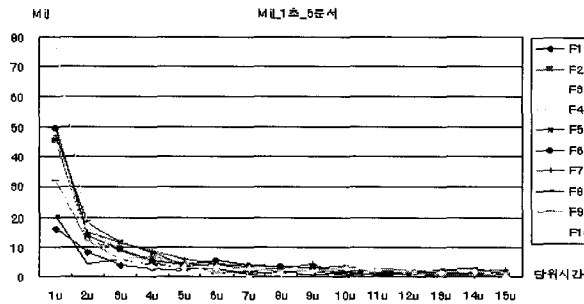
본 연구에 사용된 로그 파일은 NLANR[16]의 2001년 7

월 25일 데이터 중 요청이 가장 많았던 10시부터 11시까지의 데이터와 NASA의 1995년 7월 1일의 데이터 중의 일부분이다. 이 로그 파일들의 특성은 다음과 같다.

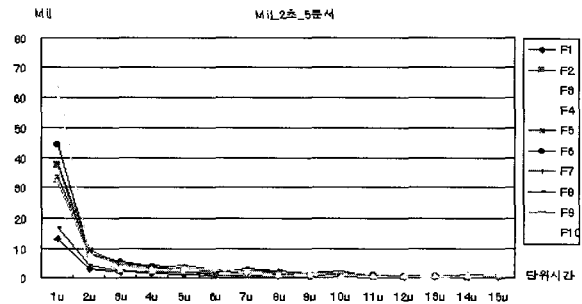
로그 파일	총 요청수	단일 요청 문서수
NLANR site RTP	60,299	34,204
NASA	60,000	1,482

Mahanti[3]의 측정도구는 스택을 기반으로 하였기 때문에 워크로드가 변하여도 T_{ij} 값에는 영향을 미치지 않는다. 즉, 시간적으로 요청수가 많아지거나 적어져도 요청 순서가 같으면 T_{ij} 값은 항상 일정하다. 그러나 임시지역성은 시간적으로 가까운 미래에 재 참조되는 것을 의미하기 때문에 워크로드와 밀접한 관련이 있다. 이러한 워크로드의 변화와 M_{iu} 값과의 관계를 알아보기 위해 워크로드를 변화시켜 실험하였다. 본 연구에서는 워크로드를 변화시키는 효과를 얻기 위해 단위시간을 변화시켰다. (그림 1)은 단위시간을 1초, 2초, 4초, 8초, 16초로 변화시키면서 측정된 M_{iu} 값을 그래프로 나타낸 것이다. 그래프의 가로축은 상호참조 단위시간이며, 세로축은 5개 문서의 M_{iu} 값들의 평균이다. F_1 은 문서 D_1 부터 D_5 까지, F_2 는 D_6 부터 D_{10} 까지, F_3 은 계속해서 그 다음 5개의 문서를 나타낸다.

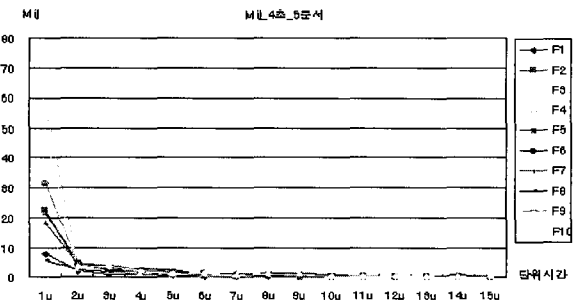
(그림 1)(a)는 단위시간을 1초, (그림 1)(b)는 2초, (그림 1)(c)는 4초, (그림 1)(d)는 8초로 하여 계산한 NLANR 로그 데이터들에 대한 M_{iu} 값의 그래프이다. 전체적으로 단위시간이 길어졌을 때, 즉 워크로드가 커졌을 때 M_{iu} 가 작아지는 현상을 나타내고 있다. 그러나 단위시간이 길어짐(위



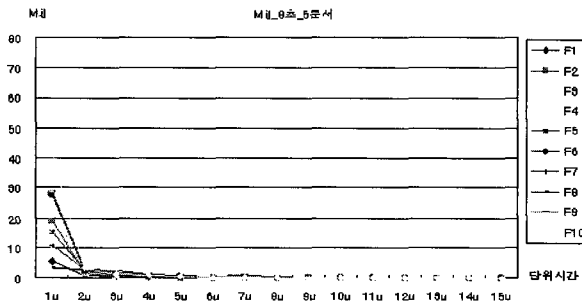
(a)



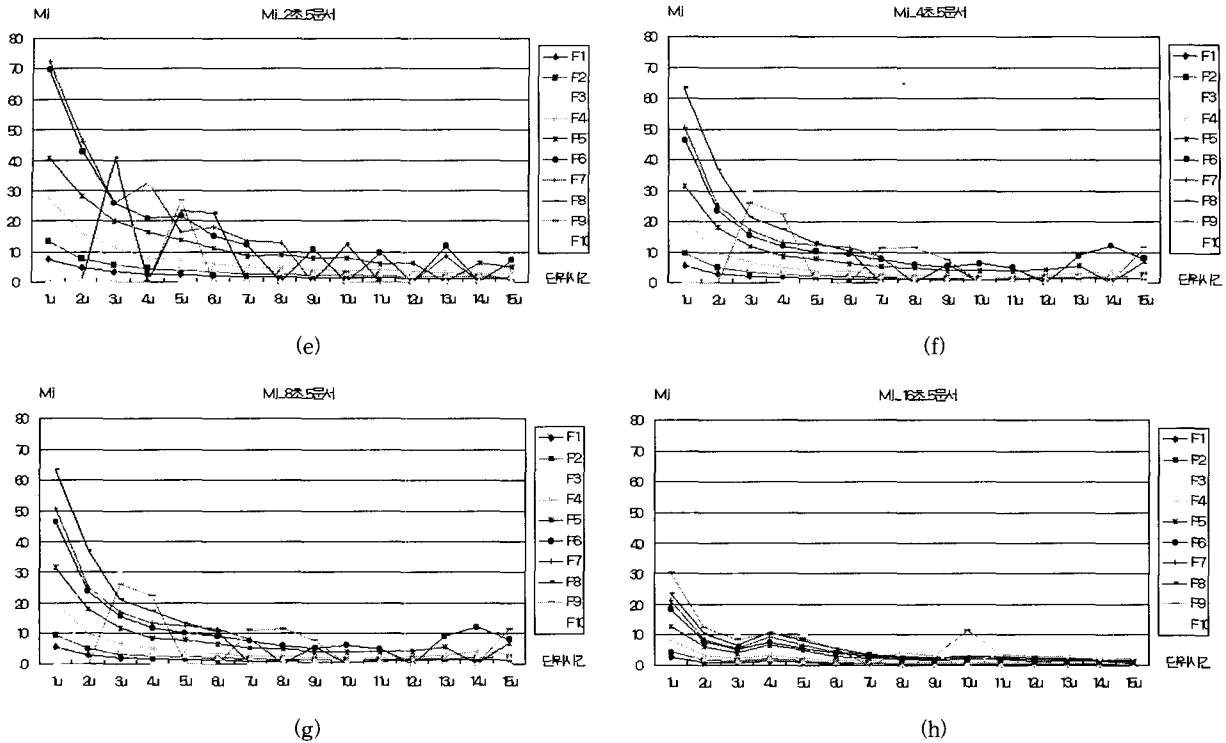
(b)



(c)



(d)



(그림 1) 단위시간 변화에 따른 M_{iu} 값

크로드가 커짐에 따라 상호참조 단위시간을 짧게 했을 때의 M_{iu} 의 변화가 상호참조 단위시간을 길게 했을 때의 M_{iu} 의 변화보다 작다라는 사실을 알 수 있다. 이는 곧 M_{iu} 값은 T_{ij} 와는 달리 워크로드와 밀접한 관계가 있다는 것을 의미한다. NASA 로그 데이터들에 대한 M_{iu} 값의 그래프인 (그림 1)(e), (그림 1)(f), (그림 1)(g), (그림 1)(h)도 유사한 경향을 보인다. 따라서 이후부터는 NLANR 로그 데이터에 대한 실험결과만 논의하도록 한다.

4.2 선호도와 M_{iu} 값과의 관계

문서의 선호도와 M_{iu} 값과의 관계를 알아보기 위해 요청수가 많은 10개의 문서와 요청수가 비교적 낮은 문서 10개에 대하여 실험하였다. 단위시간을 2초로 하여 M_{iu} 값을 계

산한 결과는 <표 4>, <표 5>와 같다.

<표 4>에서 일반적으로 상호참조 단위시간이 작을 때에는 M_{iu} 값이 비교적 큰 것을 알 수 있고, <표 5>에서는 요청수가 적은 문서들의 M_{iu} 값은 상호참조 단위시간이 작아도 많은 부분에서 0을 나타내고 있음을 알 수 있다. 따라서 문서의 선호도와 M_{iu} 는 밀접한 관계가 있다는 것을 알 수 있다. 그러나 요청수가 매우 적을 때에는 M_{iu} 값으로 임시 지역성을 측정하는데 한계가 있다. 예를 들어, D_{153} 의 M_{iu} 값은 다른 것에 비해 큰 값을 가진다. 이는 D_{153} 의 총 요청수가 31인데, 이중 1u 내에 재 요청된 수가 8이고, 1u 내에 재 요청되었을 때 그 사이에 들어온 D_{153} 을 포함한 요청 횟수는 224이다. 이 값들은 전체 요청수 60,299에 비하여 매우 작기 때문에 M_{iu} 값을 계산할 때에 작은 값으로 나누는

<표 4> 요청수 상위 10개 문서들의 M_{iu} 값

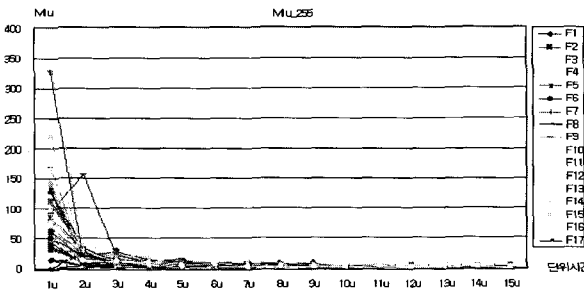
문서	1u	2u	3u	4u	5u	6u	7u	8u	9u	10u	11u	12u	13u	14u	15u	요청수
D_1	9.10	1.61	1.04	0.66	0.56	0.54	0.44	0.33	0.31	0.24	0.24	0.20	0.18	0.17	0.14	746
D_2	10.53	2.22	1.22	0.98	0.84	0.63	0.47	0.43	0.29	0.37	0.30	0.26	0.21	0.25	0.24	570
D_3	12.54	3.52	1.99	1.49	1.12	0.94	0.84	0.93	1.06	0.00	0.65	0.47	0.65	0.00	0.00	359
D_4	15.36	4.80	2.57	2.04	2.28	1.24	1.67	1.02	0.76	0.86	0.60	0.55	0.51	0.62	0.54	245
D_5	18.29	4.55	3.61	2.83	2.06	1.54	1.33	1.18	1.08	0.97	0.83	0.82	0.76	0.65	0.57	206
D_6	14.42	7.28	3.25	3.53	2.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	201
D_7	58.12	11.70	0.00	3.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	169
D_8	25.76	5.94	4.09	3.53	2.32	1.89	1.60	1.40	1.48	0.00	0.77	0.90	1.18	0.85	0.75	166
D_9	41.84	6.50	4.14	3.96	3.32	1.58	0.00	1.13	0.90	0.00	0.00	0.00	0.00	0.00	0.00	158
D_{10}	48.01	6.80	0.00	3.40	2.48	2.12	1.78	1.57	1.48	1.34	1.07	0.89	0.92	0.75	0.71	157

<표 5> 요청수 하위 10개 문서들의 M_{iu} 값

문서	1u	2u	3u	4u	5u	6u	7u	8u	9u	10u	11u	12u	13u	14u	15u	요청수
D_{151}	0.00	0.00	18.15	17.67	0.00	14.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	32
D_{152}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31
D_{153}	67.07	25.76	17.68	16.77	12.55	14.10	0.00	8.21	6.22	0.00	6.08	0.00	4.60	0.00	3.59	31
D_{154}	0.00	0.00	0.00	0.00	0.00	0.00	8.09	7.85	6.18	0.00	0.00	0.00	0.00	0.00	0.00	31
D_{155}	0.00	0.00	0.00	0.00	0.00	0.00	8.53	7.63	7.18	5.86	0.00	0.00	0.00	0.00	0.00	31
D_{156}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30
D_{157}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.73	0.00	6.18	0.00	0.00	4.19	4.07	0.00	30
D_{158}	0.00	22.84	0.00	0.00	0.00	0.00	0.00	9.26	0.00	5.60	4.47	0.00	0.00	0.00	3.62	30
D_{159}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.67	0.00	6.18	0.00	0.00	0.00	3.90	0.00	30
D_{160}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30

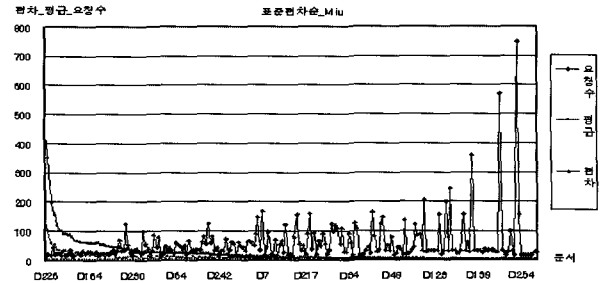
과정에서 일어나는 현상이다.

단일 요청 문서는 34,204개로 많기 때문에 이중 요청수가 많은 순서로 255개의 문서에 대한 M_{iu} 값을 계산하여 그래프로 그려보면 (그림 2)와 같다. 그래프의 가로축은 상호참조 단위시간을 나타내며, 세로축은 5개 문서의 M_{iu} 값들의 평균값을 나타낸다. 그림에서 상호참조 단위시간이 적을 때에는 M_{iu} 값이 비교적 크게 나타나다가 상호참조 단위시간이 길 때에는 M_{iu} 값이 작아지는 현상을 확인할 수 있다. 그래프에서 일부 문서의 M_{iu} 값이 매우 큰 것들이 나타나는데, 이는 해당 문서의 요청수가 전체 요청수에 비해 상대적으로 매우 적기 때문에 작은 수로 나누는 과정에서 나타나는 현상이다.



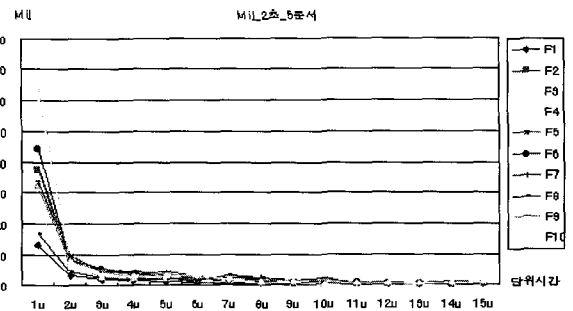
(그림 2) 요청수 상위 255개 문서들의 M_{iu} 값

(그림 3)은 M_{iu} 값의 평균과 표준편차, 그리고 문서의 요청수와의 관계를 보여주는 그래프이다. 그래프의 가로축은 M_{iu} 의 표준편차가 큰 것에서 작은 순서로 나열한 문서이며, 세로축은 M_{iu} 의 평균과 표준편차, 문서의 요청수를 나타낸다. 그림에서 M_{iu} 의 평균은 표준편차가 매우 큰 경우를 제외하고는 커다란 변화가 없다. 그러나 M_{iu} 의 표준편차는 원점에 가까울수록, 즉 요청수가 0에 가까워지면 커지는 현상을 보여주고 있다. 이는 요청수가 매우 적을 때 M_{iu} 의 표준편차가 큰 것을 의미하고, 요청수가 비교적 많을 때 M_{iu} 값의 표준편차가 작다는 것을 의미한다. 따라서 요청수가 매우 적은 경우를 제외하면 표준편차가 작다라는 것을 알 수 있다.

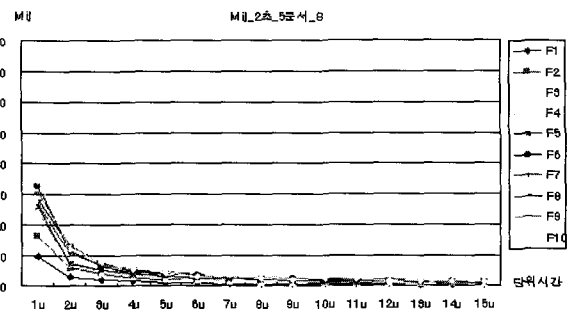


(그림 3) 표준편차 순서의 M_{iu} 값

4.3 단기간 임시지역성과 M_{iu} 값과의 관계



(a)



(b)

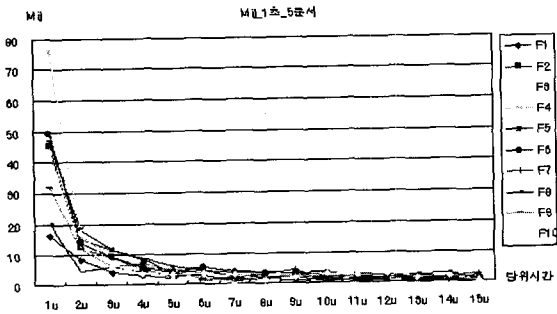
(그림 4) 요청순서에 따른 M_{iu} 값과 무작위순서에 따른 M_{iu} 값의 비교(u = 2초)

임시지역성은 문서의 선호도에 의한 장기간 임시지역성

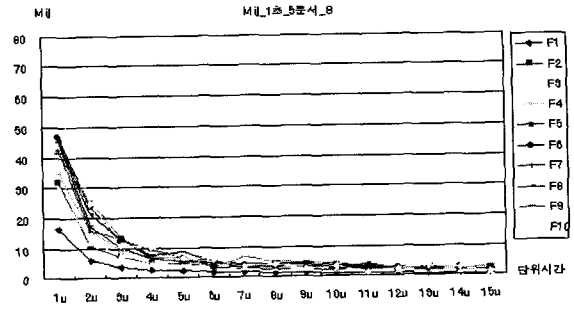
과 문서 참조의 상호관계에 의한 단기간 임시지역성이 있다. 이중 단기간 임시지역성과 M_{iu} 값과의 관계를 알아보기 위해 요청수가 많은 50개 문서에 대해 실험을 하였다.

이를 좀더 자세히 알아보기 위해 단위시간을 1, 2, 4, 8초

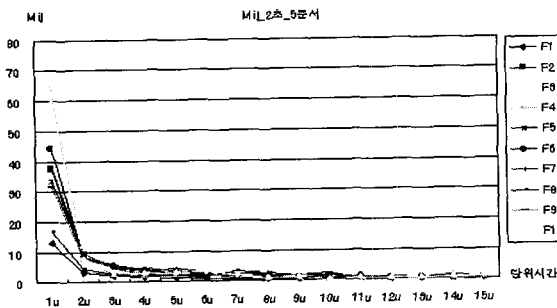
로 하여 계산한 M_{iu} 값과 이들을 무작위순서로 하여 계산한 M_{iu} 값을 그래프로 나타내면 (그림 5)와 같다. 그림의 왼쪽 (그림 5(a), (그림 5(c), (그림 5(e), (그림 5(g))는 요청흐름에 대한 M_{iu} 값을 나타내고, 오른쪽 (그림 5(b), (그



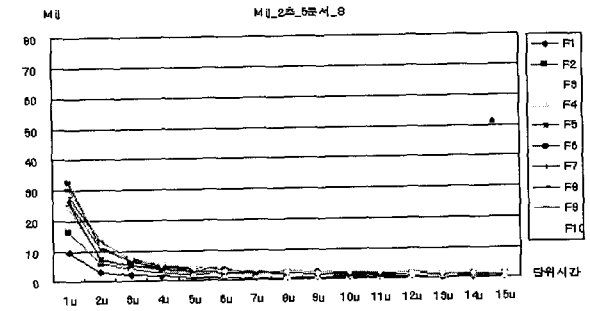
(a)



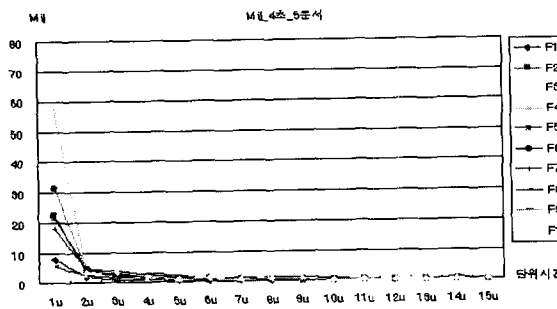
(b)



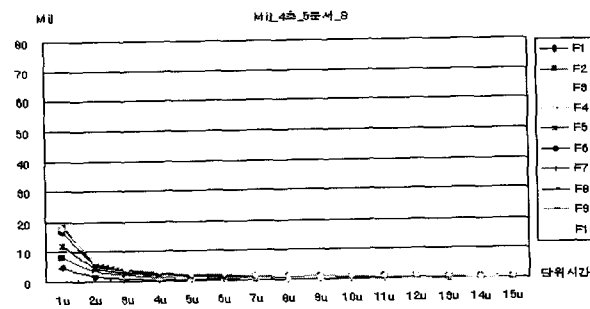
(c)



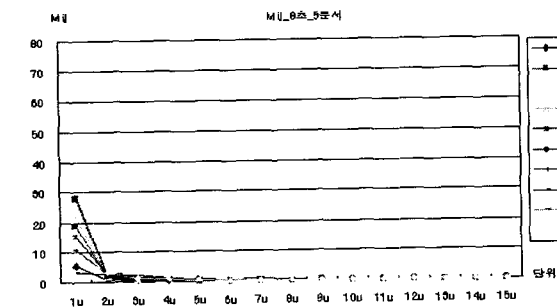
(d)



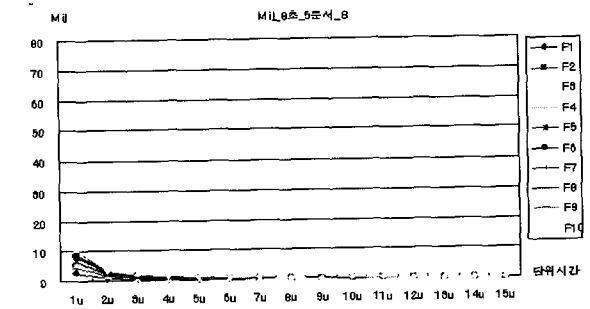
(e)



(f)



(g)



(h)

(그림 5) 단위시간 변화별 요청순서 및 무작위순서에 따른 M_{iu} 값의 비교

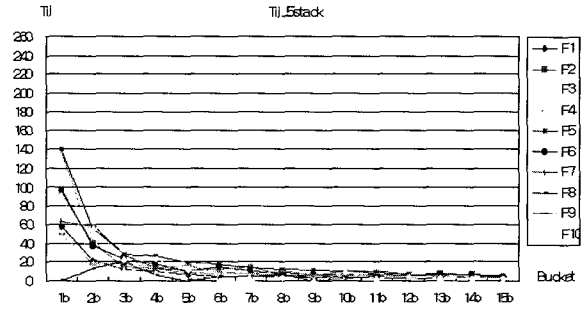
림 5(d), (그림 5(f), (그림 5(h)는 무작위순서에 따른 M_{iu} 값을 나타내는 그래프이다.

단위시간을 짧게 하면 요청수가 적게 되고, 단위시간을 크게 하면 요청 수가 많게 되는 효과가 있다. 요청흐름에 대한 M_{iu} 값은 단위시간이 커짐에 따라 조금씩 작아지는 반면, 무작위순서에 대한 M_{iu} 값은 단위시간이 커짐에 따라 현저히 작아지는 것을 보여주고 있다. 즉, 단위시간을 작게 했을 때에는 요청순서에 의한 M_{iu} 값과 무작위순서에 의한 M_{iu} 값과의 차이가 적다. 이는 단기간 임시지역성이 낮다는 것을 의미한다. 단위시간을 크게 했을 때에는 요청순위에 의한 M_{iu} 값과 무작위순서에 의한 M_{iu} 값과의 차이가 크다. 이는 단기간 임시지역성이 높은 것을 의미한다. 즉, 상호참조시간을 고려한 단기간 임시지역성 측정값은 워크로드에 따라 달라진다는 것을 알 수 있으며, 단기간 임시지역성과 M_{iu} 값은 매우 밀접한 관계가 있다는 것을 보여준다.

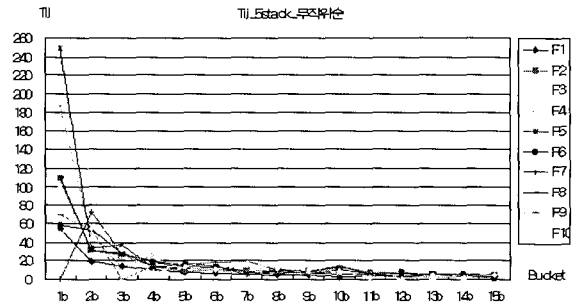
앞서 사용된 로그 파일을 이용하여 스택기반의 단기간 임시지역성 측정도구인 T_{ij} 값을 계산하였다. Mahanti[3]에서의 실험과 같이 5개의 stack 원소를 1 bucket으로 하여 계산하였고, 요청수가 많은 10개의 문서에 대한 T_{ij} 값을 보면 다음 <표 6>과 같다.

<표 6>을 좀더 쉽게 파악하기 위해 요청수가 많은 50개의 문서에 대해 T_{ij} 값을 계산하였다. 그리고 단기간 임시지역성이 정확히 측정되었는지를 알아보기 위하여 요청흐름을 난수를 이용해서 무작위순서로 나열한 다음, T_{ij} 값을 계산한 후 그래프를 그려 비교한 것이 (그림 6)이다. (그림 6)(a)는 요청흐름에 따른 T_{ij} 값이며, (그림 6)(b)는 무작위순서로 했을 때의 T_{ij} 값이다. 그림에서 요청흐름에 따른 T_{ij} 값과 무작위순서에 따른 T_{ij} 값은 서로 큰 차이가 없는 것을 알 수 있다. 일부 문서는 오히려 무작위순서로 했을 때의 T_{ij} 값이 요청순위에 따른 T_{ij} 값보다 더 크게 나타나는 경향이 있다. 이는 T_{ij} 가 M_{iu} 보다 단기간 임시지역성과 관계가 적다는 것을 의미한다. 즉, T_{ij} 는 상호참조시간을 반영하지 않고 스택을 기반으로 하기 때문에 단기간 임시지

역성 측정에 부적합하다는 것을 의미한다. 반면, M_{iu} 는 요청흐름에 대한 값과 무작위순서로 하였을 때의 값이 차이가 크게 나타났다. 이는 M_{iu} 가 T_{ij} 보다 단기간 임시지역성을 보다 민감하게 반영하고 있음을 의미한다.



(a)



(b)

(그림 6) 요청순서에 따른 T_{ij} 값과 무작위순서에 따른 T_{ij} 값의 비교

4.4 캐싱 전략과 M_{iu} 값과의 관계

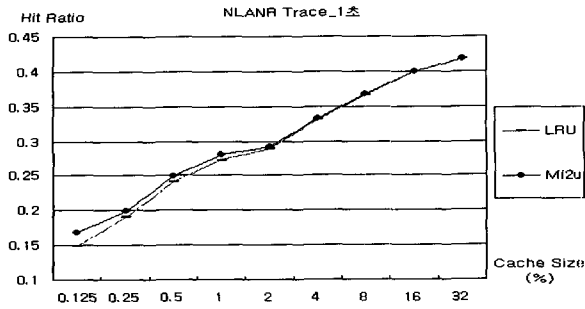
상호참조시간을 고려한 단기간 임시지역성이 실제 캐싱 전략에 어떠한 영향을 미치는지 알아보기 위해 문서 적중률 실험을 하였다. M_{iu} 값을 이용한 캐싱 전략은 캐시에 들어간 문서를 제거할 때, 문서 참조시 실시간으로 계산한 M_{iu} 값을 이용하는 것이다. 본 실험에서는 M_{iu} 값이 낮은

<표 6> 요청수 상위 10개 문서들의 T_{ij} 값

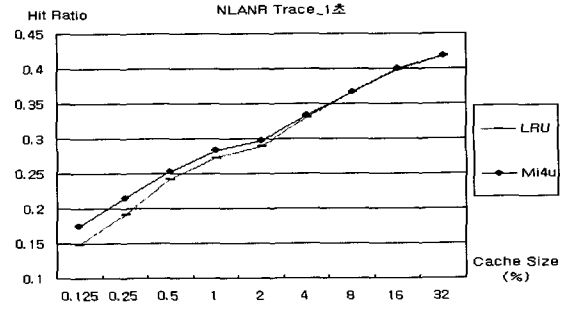
문서	1b	2b	3b	4b	5b	6b	7b	8b	9b	10b	11b	12b	13b	14b	15b	요청수
D_1	40.49	8.58	5.59	4.10	3.11	2.68	2.17	1.93	1.73	1.56	1.41	1.22	1.15	1.08	1.02	746
D_2	33.08	11.44	7.48	5.43	4.50	3.38	3.07	2.62	2.40	2.07	1.89	0.00	1.50	1.47	1.27	570
D_3	42.75	17.81	11.72	8.84	7.00	5.55	3.60	4.10	2.82	3.26	3.00	2.56	2.38	2.17	1.93	359
D_4	73.55	25.34	16.24	11.55	9.09	7.46	6.72	5.95	5.18	4.64	4.17	4.00	3.42	3.00	3.08	245
D_5	97.57	41.82	18.29	15.82	11.05	0.00	7.91	6.74	5.61	5.91	5.32	4.70	4.29	3.91	3.70	206
D_6	64.28	29.76	20.89	15.24	11.79	9.83	8.22	7.64	6.28	5.04	0.00	4.00	4.26	0.00	3.80	201
D_7	117.63	39.51	26.04	18.07	0.00	0.00	9.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	169
D_8	96.87	36.32	0.00	10.09	14.53	12.31	8.45	8.86	7.26	6.49	6.49	5.26	5.24	5.14	4.74	166
D_9	199.12	39.32	26.28	18.17	16.13	13.63	10.46	8.67	0.00	7.27	6.88	6.06	5.30	5.02	4.68	158
D_{10}	0.00	54.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.92	0.00	0.00	0.00	0.00	0.00	157

것부터 제거하였다. 그리고 스택을 기반으로 하는 캐싱 전략인 LRU 방식을 적용했을 때의 문서 적중률과 비교 실험하였다. 실험한 결과 그래프는 (그림 7)과 같다. 그래프의 가로축은 단일 문서 전체 크기에 대한 퍼센트를 나타내고, 세로축은 문서 적중률을 나타낸다.

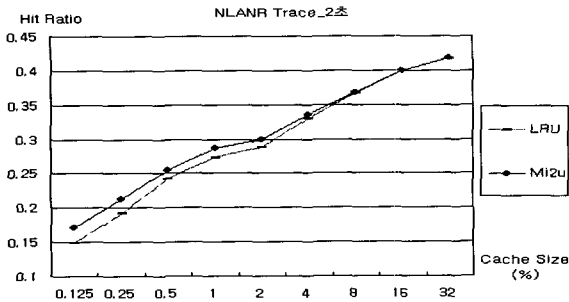
(그림 7)(a)는 단위시간이 1초이며, (그림 7)(b)는 2초, (그림 7)(c)는 4초, (그림 7)(d)는 8초로 하여 실험하였다. 또, (그림 7)(a)-1, (그림 7)(b)-1, (그림 7)(c)-1, (그림 7)(d)-1은 해당 단위시간의 2u에서 측정된 단기간 임시지역성을 캐싱 방식에 적용하였으며, (그림 7)(a)-2, (그림 7)



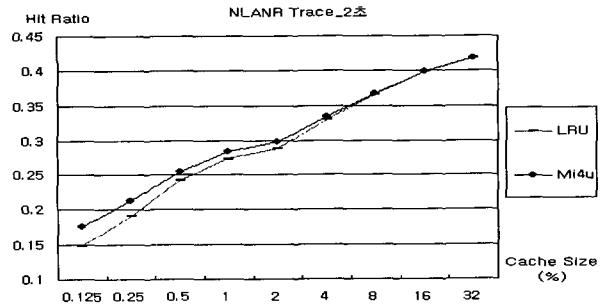
(a)-1



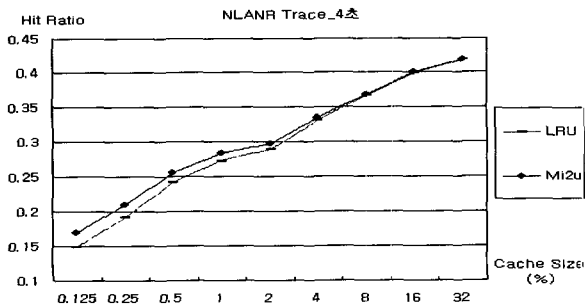
(a)-2



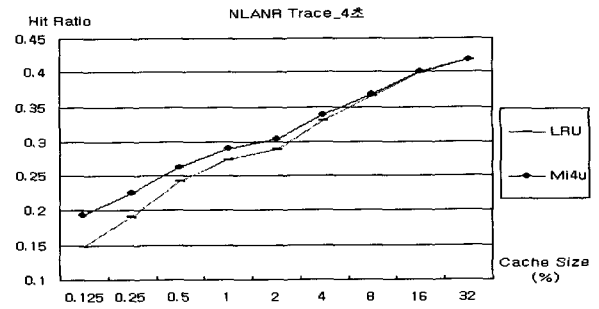
(b)-1



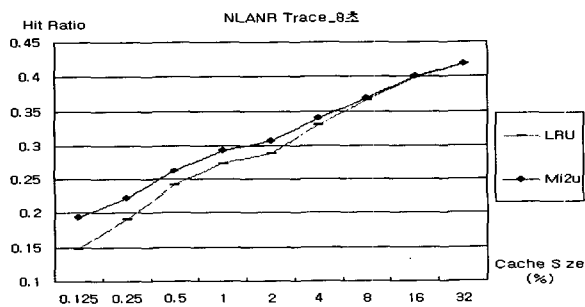
(b)-2



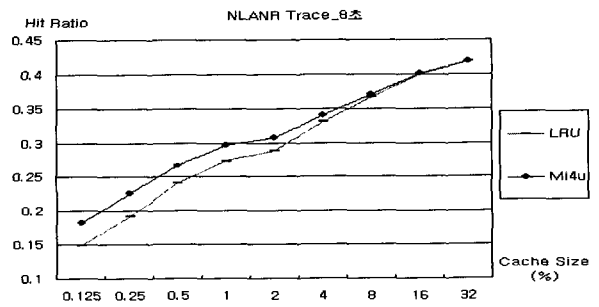
(c)-1



(c)-2



(d)-1



(d)-2

(그림 7) M_u 값으로 한 캐싱 방식과 LRU 방식과의 문서적중률 비교

(b)-2, (그림 7)(c)-2, (그림 7)(d)-2는 해당 단위시간의 4u에서 측정된 단기간 임시지역성을 캐싱 방식에 적용한 그래프이다. 단위시간과 그 단위시간의 간격에 따라 문서 적중률이 다소 변화가 있지만 대체적으로 M_{iu} 값을 적용한 캐싱 방식의 문서 적중률이 LRU 방식의 문서 적중률보다 높게 나타났다. 이는 M_{iu} 값이 캐싱 전략에 유용한 인자라는 것을 보여준다.

이상의 실험을 통해 스택을 기반으로 한 단기간 임시지역성 측정도구보다 상호참조시간을 고려한 단기간 임시지역성 측정도구가 단기간 임시지역성을 더 정확하게 측정할 수 있다는 것을 확인할 수 있다. 이는 임시지역성은 가까운 미래에 재 참조된다는 임시지역성의 정의와도 일치한다.

5. 결 론

임시지역성은 웹의 중요한 특성 중의 하나로 캐싱전략 등 웹의 성능에 많은 영향을 끼치고 있다. 그러므로 임시지역성을 정확히 측정한다는 것은 매우 중요하다. 인기도에 의한 장기간 임시지역성은 많이 참조되는 문서들을 하루하루 비교하는 것으로 간단히 측정할 수 있다. 제안한 단기간 임시지역성 측정방법은 문서별로 임시지역성의 정도를 측정할 수 있을 뿐만 아니라, 단기간 임시지역성의 정의에 부합하는 시간적 특성까지도 반영하여 측정할 수 있다. 다만 본 연구에서 제안한 측정도구 중 최적의 단위시간에 대해서는 연구가 더 필요하다.

참 고 문 헌

[1] Virgilio Almeida, Azer Bestavros, Mark Crovella and Adriana de Oliveira, "Characterizing Reference Locality in the WWW," The IEEE Conference on Parallel and Distributed Information Systems, pp.2-4, Dec. 1996.

[2] Shudong Jin and Azer Bestavros, "Temporal Locality in Web Request Streams Sources, Characteristics, and Caching Implementation," In Proceedings SIGMERICS '2000, pp.4-12, 2000.

[3] Anirban Mahanti, Derek Eager and Carey Williamson, "Temporal Locality and its Impact on Web Proxy Cache Performance," Performance Evaluation Journal : Special Issue on Internet Performance Modeling, 42 (2/3), pp.2-9, Sep., 2000.

[4] Pei Cao and Sandy Irani, "Cost-Aware WWW Proxy Caching Algorithm," Proceedings of the 1997 USENIX Symposium in Internet Technology and Systems, pp.4-5, Dec., 1997.

[5] Lee Breslau, Pei Cao, Li Fan, Gramham Phillips, and Scott Shenker, "Web Caching and Zipf-like Distributions : Evidence and Implications," Proceedings of Infocom '99, pp.1-4,

April, 1999.

[6] Shudong Jin and Azer Bestavros, "Greedy Dual Web Caching Algorithm Exploiting the Two Sources of Temporal Locality in Web Request Streams," Proceedings of the 5th International Web Caching and Content Delivery Workshop, Lisbon, Portugal, pp.2-3, May, 2000.

[7] Ludmila Cherkasova and Gianfranco Ciardo, "Characterizing Temporal Locality and its Impact on Web Server Performance," Tech-Report HPL-2000-82, pp.8-11, June, 2000.

[8] Anirban Mahanti and Carey Williamson, "Web Proxy Workload Characterization," Technical Report, Department of Computer Science, University of Saskatchewan, pp.23-25, Feb., 1999.

[9] Annie P. Foong, Yu-Hen Hu, and Dennis M. Heisey, "Web Caching : Locality of References Revisited," Proceedings of the IEEE International Conference on Networks(ICON '00), p.3, September, 2000.

[10] D. Willick, D. Eager and R. Bunt, "Disk Cache Replacement Policies for Network File Servers," Proceedings of the 13th International Conference on Distributed Systems(ICDCS), Pittsburgh, PA, pp.2-3, May, 1993.

[11] E. G. Coffman and P. J. Denning, "Operation systems theory," Prentice-Hall, 1973.

[12] J. Spirn, "Distance String Models for Program Behavior," IEEE Computer, Vol.9, pp.14-20, 1976.

[13] Shudong Jin and Azer Bestavros, "Popularity-Aware Greedy Dual-Size Web Proxy Caching Algorithm," Proceedings of The 20th International Conference on Distributed Computing Systems, 2000.

[14] John Dillely and Martin Arlitt, "Improving Proxy Cache Performance : Analysis of Three Replacement Polices," IEEE Internet Computing, Vol.3, No.6, Nov./Dec., 1999.

[15] Azer Bestavros, "Demand-based Document dissemination to Reduce Traffic and Balance Load in Distributed Information Systems," Proceedings of SPDP '95 : The 7th IEEE Symposium on Parallel and Distributed Processing, San Antonio, TX, Oct., 1995.

[16] National Laboratory for Applied Network Research. ftp://rpt.nlanr.net/Traces.



김 영 일

e-mail : yikim@kerinet.re.kr

1981년 충남대학교 전자교육공학과(학사)

1992년 아주대학교 전자계산학과(석사)

1989년~1998년 경기도과학교육원, 경기도

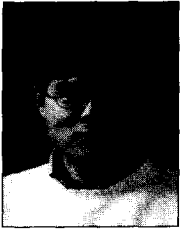
교육정보연구원

1999년~2002년 경기도교육청 장학사

2002년~현재 고색중학교 교감

1995년~현재 아주대학교 컴퓨터공학과 박사과정(수료)

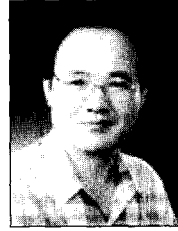
관심분야 : 운영 체제, 분산시스템, 실시간 및 멀티미디어시스템



심재홍

e-mail : jhshim@chosun.ac.kr
 1987년 서울대학교 전산학과(학사)
 1989년 아주대학교 컴퓨터공학과(석사)
 2001년 아주대학교 컴퓨터공학과(박사)
 1989년~1994년 서울시스템(주) 공학
 연구소

1999년~2000년 University of Arizona 객원연구원
 2001년~2001년 아주대학교 정보통신전문대학원 BK21 전임
 연구원
 2001년~현재 조선대학교 인터넷소프트웨어공학부 전임강사
 관심분야 : 운영 체제, 분산시스템, 실시간 및 멀티미디어시스템



최경희

e-mail : khchoi@madang.ajou.ac.kr
 1976년 서울대학교 수학교육과(학사)
 1979년 프랑스 그랑테폴 Enseiht대학
 (석사)
 1982년 프랑스 Paul Sabatier대학 정보공
 학부(박사)

1982년~현재 아주대학교 정보통신전문대학원 교수
 관심분야 : 운영 체제, 분산시스템, 실시간 및 멀티미디어시스템 등



정기현

e-mail : khchung@madang.ajou.ac.kr
 1984년 서강대학교 전자공학과(학사)
 1988년 미국 Illinois주립대 EECS(석사)
 1990년 미국 Purdue대학 전기전자공학부
 (박사)
 1991년~1992년 현대반도체 연구소

1993년~현재 아주대학교 전자공학부 교수
 관심분야 : 컴퓨터구조, VLSI 설계, 멀티미디어 및 실시간
 시스템 등