

## 웹 객체의 사용자 참조 특성 분석

### Analysis of the User's Reference Characteristics on the Web Object

고일석, 나윤지\*

충북과학기술대학 전자상거래과,  
우송대학교 컴퓨터학과\*

Il-Seok Ko, Yun-Ji Na\*

#### 요약

웹 기반의 디지털콘텐츠 전송은 웹 객체를 단위로 이루어진다. 따라서 웹 기반 콘텐츠 전송의 성능을 향상시키기 위해서는 웹 객체의 참조특성에 대한 분석이 필요하다. 본 연구에서는 웹 객체의 참조특성을 분석하였다. 또한 이 분석 결과를 통해 웹 객체의 전송 효율을 높일 수 있는 방안을 제시하였다.

#### Abstract

The unit of the digital content delivery is the web object. For the improvement of the content delivery based on the web, it is required analysis of reference characteristics of the web object. In this study, we analyze reference characteristics of the web object. Using these analysis results, we propose the method which improve the delivery performance of the web object.

## 1. 서론

웹을 기반으로 하는 시스템의 전송단위는 웹 객체이다. 따라서 웹 객체의 사용자 참조 특성은 웹 캐싱, CDN과 같은 웹 시스템의 성능에 중요한 영향을 미친다[1,2,3].

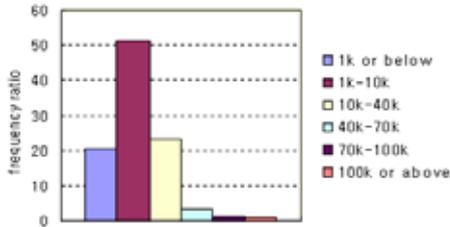
본 연구에서는 웹 객체의 참조특성을 분석하였고 또한 이 분석 결과를 통해 웹 객체의 전송 효율을 높일 수 있는 방안을 제안하였다. 분석 방법은 사용자 참조특성을 반영한 확률통계 모델을 기반으로 하고 있으며, 이를 기반으로 웹 시스템 활용에서 처리 지연 시간을 개선하는 것을 목적으로 하고 있다.

## 2. 웹 객체 참조 특성 분석

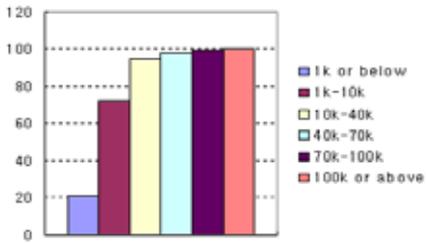
객체의 이질성 편차를 보인다.

### 2.2 참조특성 실험

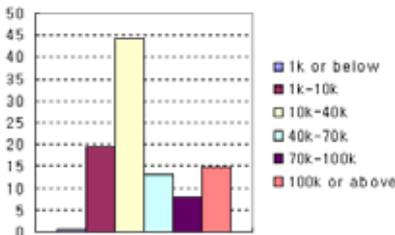
그림 1과 그림 2에서 볼 수 있듯이, 10k까지의 요청비율이 전체 참조의 약 85%에 해당한다. 하지만 그림 3과 그림 4에서 객체의 크기를 고려하면 약 21% 정도 비중을 차지하게 된다. 따라서 웹 캐시의 공간 분할에 대해 단순 크기를 비교한다면 약 2:8의 결과가 나타날 것이다. 그림 3과 그림 4의 객체의 크기를 고려한 총전송량은 웹 시스템의 성능에 많은 영향을 미친다.



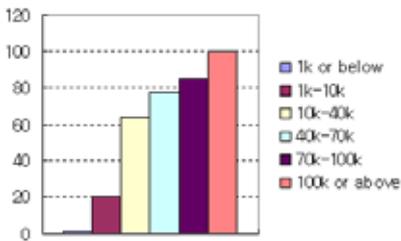
▶▶ 그림 1. Reference Frequency ratio(%)



▶▶ 그림 2. Reference frequency accumulation ratio(%)



▶▶ 그림 3. Total transmission quantity ratio(%)

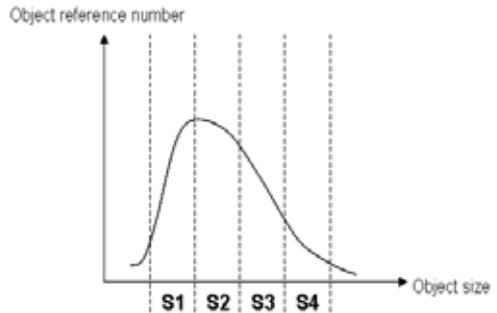


▶▶ 그림 4. Total transmission quantity accumulation ratio(%)

### III. 분석과 성능향상 방안

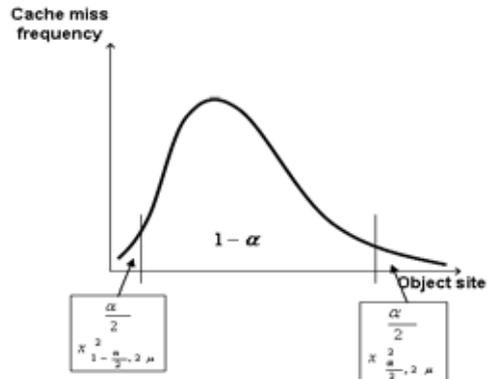
#### 3.1 객체 참조 빈도수 특성

그림 5는 객체 크기 VS. 객체 참조 빈도수를 나타낸 그래프이다. 크기가 작은 객체의 참조 빈도수가 높다. 이러한 관점에서 보면, 크기가 작은 객체를 캐시에 많이 저장하는 것이 객체적중률을 높일 수 있다. 전통적인 대부분의 웹 캐싱 기법들은 객체적중률을 높이기 위해 노력하고 있다.



▶▶ 그림 5. 객체 크기 VS. 객체 참조 빈도수

#### 3.2 Cache Miss Frequency 분포 특성



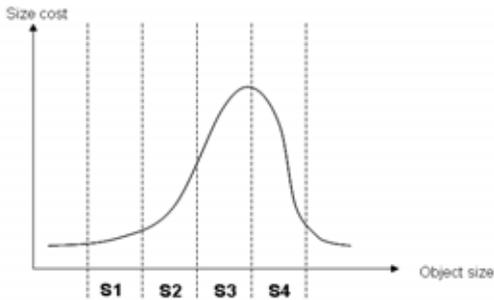
▶▶ 그림 6. Cache Miss Frequency 분포

$$P\left\{x^2_{1-\frac{\alpha}{2}, 2\mu} \leq \frac{2r\hat{\theta}}{\theta} \leq x^2_{\frac{\alpha}{2}, 2\mu}\right\}$$

$$P\left\{x^2_{1-\frac{\alpha}{2}, 2\mu} \leq x^2 \leq x^2_{\frac{\alpha}{2}, 2\mu}\right\} = 1 - \alpha$$

### 3.3 웹 캐시 비용 측면의 특성

그림 7은 객체크기 VS. 캐시비용을 나타낸 그래프이다. 크기가 큰 객체가 크기비용이 높다. 크기 비용의 증가는 네트워크 트래픽을 증가시키며, 결국 비용의 절감과, 네트워크 트래픽 감소를 위해서는 크기가 큰 객체를 캐시에서 많이 저장하는 것이 유리하다. 이 결과는 객체 참조 빈도수 특성 분석의 결과와는 반대의 관점이다. 따라서 객체 참조 빈도수 특성 분석 결과와 웹 캐시 비용 측면의 특성 분석과는 객체 적중률 VS. 비용의 트레이드오프가 발생한다.

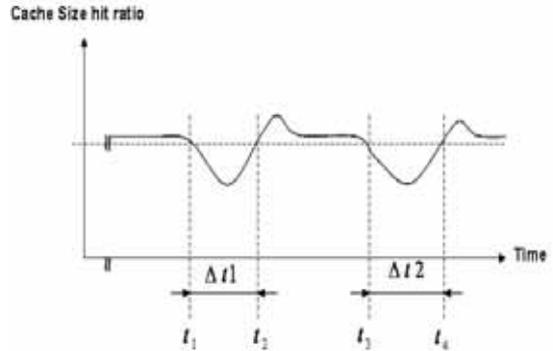


▶▶ 그림 7. 객체크기 VS. 캐시비용

### 3.4 시간 흐름에 따른 캐시 크기 적중률의 변화

그림 8은 시간의 흐름에 따른 캐시 크기 적중률의 변화를 나타낸 그래프이다. 여기에서 나타난 캐시 크기 적중률은 객체의 크기를 고려한 적중률이다. 그래프에서 t1시점과 t3시점에 적중률의 급격한 하락이 발생하여 평균값 이하로 떨어지며, t2시점과 t4시점에서 적중률의 상승에 따른 평균값 이상을 유지하게 된다. 이러한 현상이 발생하는 원인은 주로 다음과 같은 사용자의 서핑 형태 변화 특성에 따른 것이다.

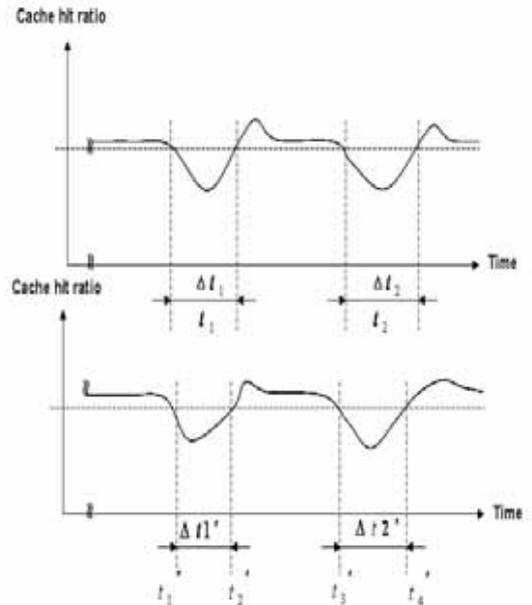
- 사용자 선호도의 변화
- 사용자 웹 서핑 형태의 변화
- 사용자의 변화 : 기존사용자 사용 종료, 새로운 사용자 사용 개시



▶▶ 그림 8. 시간 흐름에 따른 변화

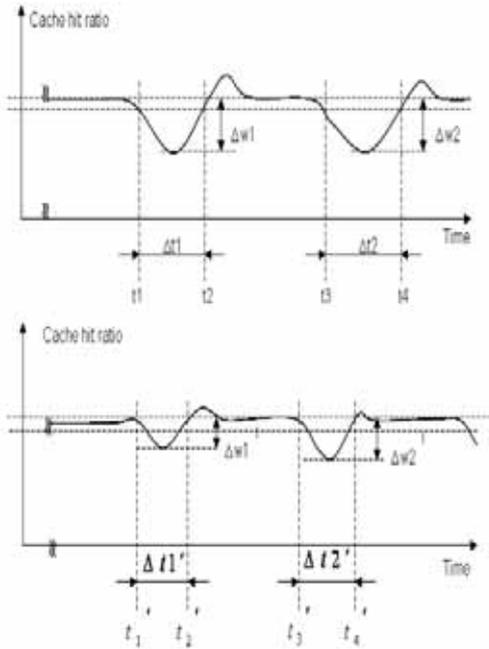
### 3.5 성능 향상 방안

①  $\Delta t1 \rightarrow \Delta t1'$ ,  $\Delta t2 \rightarrow \Delta t2'$ 의 감소



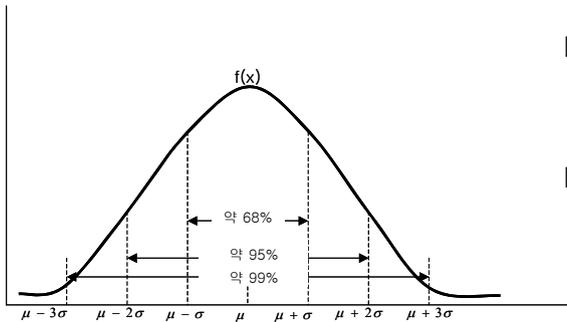
▶▶ 그림 9.  $\Delta t1 \rightarrow \Delta t1'$ ,  $\Delta t2 \rightarrow \Delta t2'$ 의 감소

②  $\Delta t1 \rightarrow \Delta t1'$ ,  $\Delta t2 \rightarrow \Delta t2'$ 의 감소 &  $\Delta w1 \rightarrow \Delta w1'$ ,  $\Delta w2 \rightarrow \Delta w2'$ 의 감소



▶▶ 그림 10.  $\Delta t1 \rightarrow \Delta t1'$ ,  $\Delta t2 \rightarrow \Delta t2'$  의 감소 &  $\Delta w1 \rightarrow \Delta w1'$ ,  $\Delta w2 \rightarrow \Delta w2'$  의 감소

③ 정규분포 활용



▶▶ 그림 11. 정규분포 특성

$\pi$ =원주율,  $e$ =상수,  $\mu$ =평균,  $\sigma$ =표준편차

$\pi$ =원주율,  $e$ =상수,  $\mu$ =평균,  $\sigma$ =표준편차 그리고 평균  $\mu$ , 분산  $\sigma^2$ 의 정규분포  $N$ 을 다음과 같이 나타낸다.

$$N(\mu, \sigma^2)$$

정규분포 공식

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

IV. 결론

본 연구에서는 웹 객체의 참조특성을 분석하였고 또한 이 분석 결과를 통해 웹 객체의 전송 효율을 높일 수 있는 방안을 제시하였다. 향후 연구로는 제안 시스템의 성능에 대한 평가를 통해 제안 기법에 대한 검증이 요구된다. 또한 더욱 다양한 웹 객체의 참조 특성 분석을 통해 동적으로 변하는 환경에서 웹 객체의 참조특성을 반영할 수 있는 방안에 대한 연구가 필요하다.

■ 참고문헌 ■

- [1] G. Barish, K. Obraczka, World Wide Web Caching: Trends and Techniques. IEEE Communications, Internet Technology Series, May 2000.
- [2] H. Bahn, S. Noh, S. L. Min, and K. Koh, "Efficient Replacement of Nonuniform Objects in Web Caches," IEEE Computer, Vol.35, No.6, pp.65-73, June 2002.
- [3] Annie P Foong, Yu-Hen Hu, Dennis M Heisey, "Web Caching: Locality of References Revisited," IEEE International Conference on Networks (ICON'00), Singapore, September 05-08, 2000.