

신뢰성 있는 분산 도메인 객체 서비스를 위한 적응형 퍼지 드롭 관리기

Adaptive Fuzzy Drop Manager for Service of Reliable Distribution Application Domain Objects

정택원 · 이종득*

TaegWon Jeong, ChongDeuk Lee*

전북대학교 응용시스템공학부 전자통신공학전공

요 약

최근에 사용자 위주의 서비스 요구가 증가됨에 따라 분산 도메인의 객체 정보를 서비스하기 위한 여러 기법들이 제안되고 있다. 본 논문에서는 이와 같은 요구에 따라 신뢰성 있는 분산 도메인 객체를 서비스하기 위한 퍼지 드롭 관리기를 제안하였다. 제안된 시스템은 사본 변수 c_{ij} 와 접근행렬 Z 를 이용하여 도메인에 접근할 수 있도록 하였으며, 도메인 내에서는 인트라 도메인 퍼지 적합도를 사상시켜 퍼지집합 사이의 관계성을 적용하여 객체들의 참조 관련성을 평가하였다. 그리고 질의를 확장하기 위해 도메인의 객체들에 대해서는 알파 레벨을 적용하여 참조 관련성에 따라 알파 레벨 집합을 생성하였다. 그 결과 우리의 기법이 다른 기법들에 비해서 성능평가가 향상됨을 알 수 있었다.

키워드 : 퍼지드롭관리기, 사본변수, 접근행렬, 퍼지적합도, 알파레벨

Abstract

A lot of methods are proposed to provide services for object informations in distributed domain to satisfy the recent increase of user-centered services. This paper proposed a method called fuzzy drop manager for the service of reliable distribution application domain objects. The proposed system accesses the domain using replica parameter c_{ij} and access matrix Z , and evaluates the reference relatedness inside the domain using the relatedness, given by the mapping of intra-domain fuzzy relevance, between fuzzy sets. Objects in the domain generated an α -level set according to the reference relatedness obtained by applying α -level to extend queries. Simulation results showed that the proposed method has better performance than the other methods.

Key Words : Fuzzy Drop Manager, Replica Parameter, Access Matrix, Fuzzy Relevance, α -level

1. 서 론

최근에 초고속 인터넷 및 모바일 서비스의 활성화로 인하여 통신 기기들을 이용한 정보 습득과 이용이 급속히 증가하고 있다. 또한 모바일 단말기의 성능은 점차 소형화, 계량화 되어 가고 있으며 이에 따른 서비스도 장소와 시간에 구애받지 않는 사용자 위주의 서비스로 진화하고 있다. 분산 응용 도메인 환경에서 사용자 위주의 정보 서비스가 유지되기 위해서는 끊임없는 스트리밍 서비스, 지터(jitter) 문제 최소화, 혼잡제어 등의 QoS가 보장되어야 한다[1-4]. 특히 분산 모바일 환경에서 멀티미디어 객체 서비스는 넓은 대역폭과 고 품질의 전송서비스가 요구되기 때문에 이를 위한 신뢰성 있는 정보서비스가 보장되어야 한다[5-6]. 무엇보다 사용자가 요구하는 정보는 응용 도메인 상에 분산되어 있으므로 이러한 정보를 보다 빠르고 효율적으로 스트리밍하기 위해서는 이들 정보를 신뢰성 있는 정보로 구성해야

한다. 분산 응용 도메인에서 사용자에게 제공되는 서비스는 실시간 서비스와 비실시간 서비스로 제공되고 있으며 이들의 정보는 적시 스트리밍 서비스(Timely Streaming Service)와 QoS가 보장되어야 한다[2,4]. 이러한 기능 보장을 위해 분산 도메인 서비스 환경에서 실시간으로 정제된 멀티미디어 객체를 연속적인 흐름상에서 처리 및 서비스되도록 프록시 캐싱서비스 기법이 제안되고 있다[1].

이 기법은 미디어 객체들을 전체적으로 캐싱하는 기법이 아니라 스트리밍이 발생하는 세그먼트 일부를 캐싱하는 기법이다. 그러나 이 기법은 정제되지 않는 멀티미디어 객체를 연속적인 흐름상에서 스트리밍 할 경우 객체크기, 지연 시간, 끊임 없는 스트리밍 요구 등으로 인한 서비스 객체 접근 실패, 바이트 히트율 감소, 검색 서비스율이 떨어지는 문제가 발생되고 있다. 이러한 문제점을 개선하기 위하여 세그먼트 기반의 프록시 캐싱 기법[1,7]이 제안되고 있으며 이 기법은 클라이언트에서 시작 시간을 줄이기 위한 기법으로서 네트워크 트래픽과 서버 워크 로드를 줄이기 위한 기법이다. 그러나 이 기법 또한 캐싱 시스템에서 객체 세그먼트들이 정제되지 않는 상태에서 세그먼트 들이 캐싱되기 때문에 프록시 지터와 같은 QoS 문제가 발생되고 있다. 따라

접수일자 : 2009년 1월 9일

완료일자 : 2009년 7월 22일

* 교신 저자

서 본 논문에서는 분산 도메인 환경에서 연속적인 세그먼트 객체 스트리밍 서비스가 보장되고 신뢰성 있는 사용자 주문형 서비스가 보장되도록 퍼지 드롭(Fuzzy Drop) 시스템을 제안한다. 제안된 시스템은 분산 도메인에서의 멀티미디어 응용 객체들을 추출하여 적합도를 판단하며, 적합도는 인트라 적합도(Intra Relevance)와 인터 적합도(Inter Relevance)로 구분하여 측정한다. 인트라 적합도는 참조 관련성이 높은 객체들의 신뢰성 보장을 위해 알파 레벨을 사용한다. 그리고 인터 적합도는 도메인 내에서의 객체들 간의 적합성을 측정하기 위해 사용되며, 키 참조기법 즉 KRF(Key Reference Frequency)를 사용한다. 제안된 논문에서는 적합도에 따라서 스트리밍이 수행되는 세그먼트들이 정제되며 정제된 세그먼트들은 퍼지 드롭 테이블에 의해서 관리된다. 이때 퍼지 드롭 테이블은 참조 관련성 객체 클래스와 드롭 객체 클래스를 분류한 후 동기화를 수행한다. 따라서 동기화가 수행된 객체클래스들은 스트리밍 인식이 보장되어 신뢰성 있는 스트리밍 서비스가 수행되게 된다. 본 논문의 구성은 다음과 같다. 2장에서는 관련연구에 대해서 살펴하며, 3장에서는 제안된 시스템 모델에 대해서 기술한다. 그리고 4장에서는 제안된 시스템에 대한 성능 평가에 대해서 알아보고 결론 및 향후연구에 대해서 살펴본다.

2. 관련연구

최근에 모바일 및 분산 응용환경에서 사용자 지향의 멀티미디어 응용 서비스 요구가 증가됨에 따라 효율적이고 신뢰성 있는 멀티미디어 서비스를 제공하기 위한 미러링, 캐싱, 콘텐츠 분산과 같은 여러 기법들이 제안되고 있다[5-6]. 그러나 이 기법들은 사용자들에게 콘텐츠 객체 정보를 보다 쉽게 접근할 수 있는 기능을 제공하고 있지만 서비스 효율성 면에서 신뢰성이 떨어지는 문제점이 발생되고 있다. 일반적으로 멀티미디어 분산 서비스 기법은 CMD (Centralized Multimedia Distribution) 서비스 기법[5], proxy 캐싱 서비스 기법[3,7], CDN(Content Distribution Network) 서비스 기법[5], P2P 기반의 네트워크 서비스 기법[5-6] 등이 주로 이용되고 있으며, CMD 서비스는 스트리밍과 I/O 용량을 확대하여 중앙 서버의 서비스 기능을 개선하기 위한 기법이다. 이 기법은 기존의 웹 기반 분산 서비스 기법에서 주로 활용되는 기법으로서 멀티미디어 분산 서비스의 성능에 중요한 영향을 미치는 네트워크 병목 현상이 발생하는 문제점을 가지고 있다. proxy 캐싱 서비스 기법은 자주 사용되는 콘텐츠 객체를 미리 캐싱하여 멀티미디어 분산 서비스 시스템에서 발생하는 네트워크 병목 현상을 줄이기 위한 기법이다. 이 기법에서 클라이언트들은 인터넷 대역폭을 소모하지 않고 에지 서버로부터 스트리밍 서비스를 수행한다.

CDN 기반의 서비스 기법은 분산 콘텐츠를 전송하기 위한 네트워크 서비스 기법으로서 서버 중심 서비스 기법이다. 이 기법에서 최초의 서버는 복제를 통하여 관리되며 네트워크 공간상에서 지역적 또는 원격으로 배치되게 된다. 그러나 캐싱기법과 CDN 기법은 멀티미디어 데이터를 전송하는데 있어서 데이터 끊김 현상으로 인하여 서비스 접근 효율성이 떨어지는 문제가 발생되고 있다.

이외에도 Alkami[8], Napster[9], Gnutella[10], FreeNet[11], CenterSpan[12] 등과 같은 P2P 기반의 분산

서비스 기법이 제안되고 있으며 이 기법들은 스토리지에 따른 QoS 문제가 발생되고 있다. 또한 토폴로지 인식 오버레이 기법은 분산 네트워크 환경에서 같은 네트워크 구조에 있는 이웃한 호스트와 피어(peer)들을 응용 그룹으로 클러스터링 하기 위한 기법[13]이다. 이 기법에서 응용 그룹을 구성하는 목적은 그룹 안에 있는 피어들은 서로 관련성을 가지고 있기 때문에 스트리밍 서비스가 요구될 경우 QoS를 보다 효율적으로 제공하기 위해서이다. 그러나 오버레이 기법에서 스트리밍 서비스는 네트워크 레이어에서 수행되는 스트리밍 서비스만큼 효율적이지 못하다는 문제점이 발생되고 있다. 또한 라우팅 오버헤드로 인한 오버레이 성능이 떨어지는 문제점이 발생되고 있으며 특히 오버레이가 랜덤하게 구성되면 오버레이 네트워크에서 이웃한 호스트들은 실제로 관련성이 떨어지는 문제점이 발생되고 있다.

3. 제안된 시스템 구조

본 논문에서 제안된 AFDM(Adaptive Fuzzy Drop Manager)는 그림1에서 살펴보듯이 도메인 접근을 통해서 사상 스케줄러에 의해 사용자의 객체가 서비스 될 수 있도록 클래스로 구조화를 수행한다. AFDM은 도메인 질의 확장을 통하여 알파 레벨[14]를 수행할 객체들을 결정하고 결정된 객체들에 대해서 드롭을 수행한다. 그리고 스트리밍 요청이 수행되는 드롭 객체들에 대해서 브로드캐스팅하게 된다.

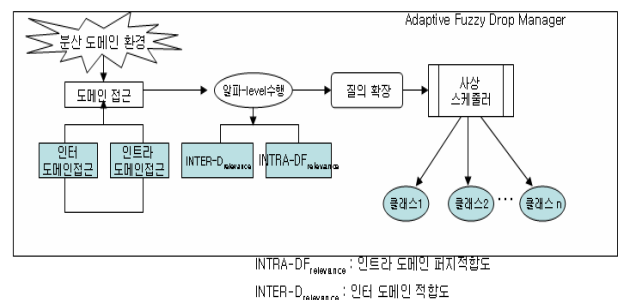


그림1. 시스템 구성도

Fig.1. The proposed system structure

그림1에서 도메인 접근은 인트라 도메인 접근과 인터 도메인 접근의 상호 작용을 통해서 수행되며 접근이 완료된 객체들은 도메인 질의 확장을 통해서 사상스케줄러로 브로드캐스트 된다. 이처럼 도메인 접근은 명시적인 요청들을 알파 레벨을 통해서 점진적으로 처리하며, 사상 스케줄러는 신뢰성 있는 서비스 객체를 생성하게 된다. 브로드캐스트가 수행된 후 필요한 서비스 객체들은 신뢰성 테이블로 구조화 된다. 도메인 질의 확장을 통하여 알파 레벨 수행과정에서 요청을 만족하는 객체를 찾지 못하면 피드백을 수행하여 요청을 만족하는 객체를 다시 브로드캐스트하게 된다.

3.1 퍼지 기초

이 절에서는 본 논문에서 적용되는 몇 가지 기본적인 퍼지 이론에 대해서 살펴본다. 퍼지 집합(Fuzzy set)이란 도메인 D에서 퍼지 집합 A는 $[0,1]$ 까지의 소속함수 μ 로 정의되며, A에서 각 아이터은 μ 의 소속값을 가지게 된다. 이때 $\phi(S)$ 를 기존의 아이터 집합 S로부터 생성된 퍼지 집합이라

하고 S의 각 아이템은 [0,1] 사이의 소속값을 가지게 되며 S를 보통집합(crisp set)이라 한다[14].

(정의1) 퍼지 관계(Fuzzy Relation): 도메인 $G \times M$ 에서의 퍼지 집합 A는 G, M 에 대해 퍼지 관계이다.

(정의2) 퍼지 교집합(Fuzzy Sets Intersection): 집합 A와 B의 퍼지 교집합은 $A \cap B$ 이고, $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$ 이다.

(정의3) 퍼지 합집합: 집합 A와 B의 퍼지 합집합은 $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$ 이다.

(정의4) 퍼지 카디널 집합: S_f 를 도메인 D에서의 퍼지 집합이라고 할 때 퍼지 카디널 집합은 $|S_f| = \sum_{x \in D} \mu(x)$ 이다. 여기서 $\mu(x)$ 는 S_f 에서의 x의 소속정도 즉 퍼지값이다.

(정의5) 퍼지 유사도 집합: 두 퍼지 집합 A와 B의 유사도(similarity)는 $\text{Sim}(A, B) = \frac{|A \cap B|}{|A \cup B|}$ 이다.

(정의6) 퍼지 subsethood 집합: 개념 클러스터 B에 대한 퍼지 집합 A의 subsethood는 $\text{subsethood}(A, B) = \frac{|A \cap B|}{|B|}$ 이다.

(정의7) 퍼지 집합의 최대-최소 조합집합(Max-Min Composition): $P(X, Y)$ 를 X, Y에 대한 퍼지 관계라 하고, $P(Y, Z)$ 를 Y, Z에 대한 퍼지 관계라 할 때 $P(X, Y)$ 와 $Q(Y, Z)$ 에 대한 Max-Min 조합 $P \cdot Q$ 는 $\mu_{P \cdot Q}(x, z) = \max_{y \in Y} \min(\mu_p(x, y), \mu_q(y, z))$, $\forall x \in X, y \in Y$ 이다.

여기서 max-min 조합은 원소 X와 Z사이의 관계성 정도를 나타낸다.

(정의8) X가 유한집합인 경우 $X = \{x_1, x_2, \dots, x_n\}$ 일 때 퍼지 집합 A를 $A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \dots + \mu_A(x_n)/x_n$ 은 $A = \sum_{i=1}^n \mu_A(x_i)/x_i$ 이다. 여기서 +는 덧셈기호가 아닌 “또는” “or”를 의미한다.

(정의9) $A = \sum_{i=1}^n \mu_A(x_i)/x_i \Rightarrow A = \{\mu_A(x_1)/x_1, \mu_A(x_2)/x_2, \dots, \mu_A(x_n)/x_n\}$ 이다.

3.2 도메인 접근

도메인 접근은 분산 응용 도메인에 분포된 도메인 객체들을 알파-레벨을 통해서 적합도를 판정하게 되며, 적합도에 따른 관련성 여부는 질의 확장을 통해서 수행된다. 도메인의 객체 참조성은 인터 도메인 접근과 인트라 도메인 접근을 통해서 이루어지며, 인터 도메인 접근은 참조 관련성 빈도(RRF : Reference Relevance Frequency)에 따라 수행된다. 그리고 인트라 도메인 접근은 객체들 간의 신뢰성을 위해 사용되며 알파 레벨에 의해 수행한다.

3.2.1 인터 도메인 적합도

인터 도메인 적합도(INTER-D_{relevance})는 도메인 간의 적합성으로서 본 논문에서는 도메인 적합도를 위해서 키 참조 빈도(KRF : Key Reference Frequency)를 씨드(seed)로 한다. 주어진 KRF에 대해서 KRF는 적합도와 비례하도록 한다. INTER-D_{relevance} 과정에서 KRF는 도메인 안에서 객체들에 대해 많은 사본들을 가질 수 있으며 이러한 사본들에 대해서 KRF를 수행한다. INTER-D_{relevance}에서 중요한 문제는 저장 공간이며, 서로 다른 도메인의 객체 저장 공간을 참조하는 객체 공간을 DRC(Domain Reference

Capacity)라 한다. 예를 들어 분산 응용 환경에서 M개의 도메인이 있다고 가정하고 도메인 모집단 D_i 는 r_i ($i=1, \dots, M$)이며, 각 도메인의 크기는 비슷하다고 가정한다. 이때 도메인 내에는 G개의 관련 그룹이 있다고 가정하고 그룹 G_i 에서 객체의 저장 공간을 S_j 라고 가정하자.

그리고 도메인 내에서 그룹과 그룹 사이의 거리 s와 t는 $d(G_s, G_t)$ 라고 하자. 인터 도메인 적합도 전략은 AFDM에 의해서 요구된 인터 도메인의 평균거리가 최소화되도록 하기 위해 서로 다른 도메인들 간의 KRF가 어느 정도인지를 탐색한다. 본 논문에서는 이와 같은 KRF를 결정하기 위하여 KRF를 위한 도메인 사본 C를 이용하며, 사본 변수 c_{ij} 를 다음과 같이 정의한다.

(정의10) $c_{ij} = \begin{cases} 1, & \text{도메인 } D_i \text{는 인터도메인에 대한 최소거리 접근 행렬을 가지고 있을} \\ 0, & \text{otherwise} \end{cases}$

따라서 인터 도메인 D_i 에 대한 접근 행렬 $Y = \{y_{ij}\}$ 를 다음과 같이 정의한다.

(정의11) $y_{ij} = k; k \neq j, \text{dist}(D_j, D_k) = \min\{\text{dist}(D_j, D_k), c_{i,k} = 1\}$ 이다.

인터 도메인 접근 행렬은 각 도메인의 분포에 따른 거리를 결정하게 되며, 이때 도메인 내에서의 그룹 간 접근 행렬 Z는 도메인 사본 C에 의해서 결정된다. 따라서 도메인 내에서의 그룹 간 거리 최소화는 다음과 같이 정의된다.

(정의12) 최소화 : $\sum_{i=1}^n \sum_{j=1}^n (r_i \times \text{dist}(G_i, G_j))$ 이며, $\sum_{i=1}^M (c_{ij}) \leq S_j, c_{ij} \in \{0,1\}, j=1,2, \dots, N$ 이다.

여기서 r은 radius로서 그룹 내에서의 거리를 결정하는 반지름이다.

그룹 내에서의 최소화는 가중치를 가진 평균 최소 거리이며, 각 그룹의 저장 능력은 각 그룹의 씨드 용량을 초과해서는 안 된다. 본 논문에서는 각 도메인에 대해서 k_i 개의 사본이 존재한다고 가정하기 때문에 K-center 문제[5]에 대해서는 고려하지 않는다.

3.2.2 인트라 도메인 퍼지 적합도

인트라 도메인 퍼지 적합도는 인터 도메인의 거리 최소화가 결정될 때 도메인 내에서의 객체의 참조 관련성을 결정하기 위한 전략이다. 본 논문에서는 인트라 도메인의 참조 관련성을 위해 알파 레벨을 사용한다. 알파레벨은 퍼지 적합도에 따라 참조 관련성을 수행하게 되며, 참조 관련성이 높은 객체들은 서비스 신뢰성을 보장받게 된다. 신뢰성 보장을 위한 알파 레벨은 도메인 퍼지 적합도(INTRA-DF_{relevance})에 의해 수행되며, 도메인 D에서의 알파 레벨을 위한 A를 X에 있어서 퍼지 집합이라 하고, $\alpha \in [0,1]$ 이라 하면, A의 α -레벨 집합을 α -cut 이라 하며 A_α 는 하며 다음과 같이 정의된다.

(정의13) $A_\alpha = \{x | \mu_A(x) \geq \alpha, x \in X, \alpha \in [0,1]\}$ 이다.

일반적으로 $a \in R$ 일 경우, $\mu_a(x) = \begin{cases} 1 & x = a \\ 0 & x \neq a \end{cases}$ 로 표현된다.

우리는 퍼지수가 정규이고 도메인에서 관련도가 매우 높은 R 상의 퍼지부분집합을 $\exists x \leftarrow R, \forall x \mu_A(x) = \alpha$ 로 표현한다. 이것은 R상의 퍼지집합의 소속 값(membership value)의 최대치가 1이며, 이는 참조 관련성이 매우 높음을 의미한다. 참조 관련성이 매우 높은 정도는 수평축에 평행한 α -cut가 $A_\alpha = [a_1^{(\alpha)}, a_2^{(\alpha)}]$ 의 성질을 포함하며, 다음과 같은 포함관계

가 성립한다.

(정의14) $(\alpha' < \alpha) \Rightarrow (a_1^{(\alpha')} \leq a_1^{(\alpha)}, a_2^{(\alpha')} \geq a_2^{(\alpha)})$ 이면, $(\alpha' < \alpha) \Rightarrow (A\alpha \subset A\alpha')$ 이다.

그리고 레벨 집합에서 퍼지집합에 의한 분해 정리는 다음과 같이 정의한다.

(정의15) $A = \bigcup_{\alpha} \alpha A\alpha$, $0 \leq \alpha < 1$ 이다. 여기서 $\alpha A\alpha$ 는 스칼라 α 와 레벨집합 $A\alpha$ 의 대수곱이며, $\alpha A\alpha$ 의 소속 함수는 $\mu_{\alpha A\alpha}(x) = \alpha \mu_{A\alpha}(x)$, $x \in X$ 이다.

따라서 도메인 D_A, D_B 가 참조 관련성을 만족하는 퍼지집합일 때 $D_{A\alpha}, D_{B\alpha}$ 및 $(D_{A\alpha} \cup D_{B\alpha})$ 에 대한 참조 관련성을 다음과 같이 정의한다.

(정의16) $(D_A \cup D_B)\alpha = \{x | \max(\mu(x)D_A, \mu(x)D_B) \geq \alpha\}$
 $= \{x | \mu D_A(x) \geq \alpha \text{ or } \mu D_B(x) \geq \alpha\}$
 $= \{x | \mu D_A(x) \geq \alpha\} \cup \{x | \mu D_B(x) \geq \alpha\}$
 $= D_{A\alpha} \cup D_{B\alpha}$

예를 들어 $X = o_1 + o_2 + o_3$ 로 하고, X 에 있어서 도메인 퍼지집합 D_A 와 D_B 를 각각

$D_A = 0.5/o_1 + 0.7/o_2 + 0.9/o_3$, $D_B = 0.4/o_1 + 0.6/o_2 + 0.8/o_3$ 라 하면 각각의 레벨 집합은

$$D_{A0.5} = \{o_1 + o_2 + o_3\}$$

$$D_{A0.7} = \{o_2 + o_3\}$$

$$D_{A0.9} = \{o_3\}$$

$$D_{B0.4} = \{o_1 + o_2 + o_3\}$$

$$D_{B0.6} = \{o_2 + o_3\}$$

$D_{B0.8} = \{o_3\}$ 에 의해서 도메인 집합이 다음과 같이 표현된다.

$$D_{A0.4} \cup D_{B0.4} = \{o_1 + o_2 + o_3\}$$

$$D_{A0.5} \cup D_{B0.5} = \{o_1 + o_2 + o_3\}$$

$$D_{A0.6} \cup D_{B0.6} = \{o_2 + o_3\}$$

$$D_{A0.7} \cup D_{B0.7} = \{o_2 + o_3\}$$

$$D_{A0.8} \cup D_{B0.8} = \{o_3\}$$

$$D_{A0.9} \cup D_{B0.9} = \{o_3\}$$
이다.

3.3 질의 확장

질의 확장은 도메인 접근을 통해서 KRF가 높은 도메인들 중에서 참조 관련성이 높은 객체들을 추론하기 위한 전략이다. 질의 확장은 도메인의 KRF를 이용하여 AFDM을 적절하게 표현할 수 있도록 질의어를 확장하고 의미적으로 연결되는 객체들에 대해서는 동일한 카테고리를 생성하여 참조 관련성이 높은 객체그룹과 드롭성이 있는 객체그룹을 적용성 있게 분류하게 된다. KRF가 수행된 도메인에서 특정 객체의 발생빈도가 높으면 높을수록 그 객체는 도메인에서 중요한 의미를 가지고 도메인의 주제를 많이 반영한다고 할 수 있다. 본 논문에서는 사본 변수 c_{ij} 와 접근행렬 Z 를 이용하여 도메인에 접근할 수 있도록 하였으며, 도메인 내에서는 인트라 도메인 퍼지 적합도를 사상시켜 퍼지집합 사이의 관계성을 적용하여 객체들의 참조 관련성을 평가하였다. 그리고 질의를 확장하기 위해 도메인의 객체들에 대해서는 알파 레벨을 적용하여 참조 관련성에 따라 호환 알파 레벨 집합을 생성하였다. 이러한 알파 레벨 집합은 사용자의 관심도 및 서비스율, 검색시간을 단축시킬 뿐만 아니라 수작업을 통한 객체 사상기법에 비해 검색기법의 단점을 개선하는 효과를 가져 오게 된다. 질의 확장을 위한 알고리즘은 다음과 같다.

Input : 도메인 및 도메인의 객체들

Output : 참조 관련성 객체들

begin

// 접근한 응용 도메인을 입력한다.

while($D_n \leq m$) begin

// D_n 은 접근되는 도메인의 수이고, m 은 전체 도메인의 수이다.

while($i < 2$) begin

(1) KRF를 이용하여 도메인을 분석한다.

(2) KRF를 위한 c_{ij} 와 Z 를 계산한다.

end

if($D_n == KRF$) then begin

(3) 인트라 도메인 퍼지 적합도를 수행한다.

(4) 알파 -레벨을 수행한다.

end

else if ($D_n \neq KRF$)

(5) next D_{n-1} 을 탐색한다.

end

(6) 임의의 α 값 이상인 객체들에 대해서 참조관련성 객체 클래스와 드롭 객체 클래스를 생성한다.

3.4 사상 스케줄러

사상 스케줄러는 질의 확장을 통해서 추론된 참조 객체들에 대해서 의미적 관계성을 수행한다. 일반적으로 분산응용도메인에서 표현되는 객체는 실행 시에 발생하는 객체들이며, 이 객체들은 도메인의 특성들과 관련성을 가진다. 도메인의 특성을 가진 객체들은 공간적, 시간적, 구조적 관계에 의해 제약을 받으며, 제약조건 C 가 $c \in C$ 이 존재 할 때 사상 스케줄러는 disjunct(\cup)와 conjunct(\cap)사상 그리고, 드롭 사상으로 구분하여 스케줄러를 수행한다.

3.4.1 disjunct사상

disjunct에 의한 사상은 일반화 관계에 따라 제약조건을 보다 일반적인 제약조건으로 사상하기 위한 방법으로서 이 방법은 분산 도메인 환경에서 시간(temporal) 및 공간(spatial)동기화를 위한 사상기법이다. 시간동기화에 의한 사상은 $M_{disjunct}^t$, 공간동기화에 의한 사상은 $M_{disjunct}^s$ 이며, disjunct에 의한 사상은 다음과 같이 정의한다.

(정의17) $M_{disjunct}^{t,s} : \forall c, c' \in C : c$ 는 $c \vee c'$ 이다.

3.4.2 conjunct사상

conjunct에 의한 사상은 보다 명세적인 드롭구조를 생성하기 위한 방법이며, 논리적, 공간적, 시간적 동기화에서 관계성이 적은 객체들을 사상하기 위해 사용된다. 논리적 동기화에 의한 사상은 $M_{conjunct}^t$, 공간적 동기화에 의한 사상은 $M_{conjunct}^s$, 시간적 동기화에 의한 사상은 $M_{conjunct}^t$ 이며, 다음과 같이 정의된다.

(정의18) $M_{conjunct}^{t,s,t} : \forall c, c' \in C : c$ 는 $c \wedge c'$ 이다.

$\forall c \in C, \emptyset$ 은 c 의 null이다.

3.4.3 드롭 사상

드롭사상은 참조 드롭 테이블에서 의미적으로 관련성이 적은 객체들을 필터링하기 위한 사상이다. 응용 도메인을 D 에서 필터링을 위한 사상은 $f_i \in F$, 필터링에 의해서 분류된 객체 그룹은 $g_i \in G$ 라 할 때, 도메인에서의 필터링을 위

한 드롭 사상($M_{(o,1) \cup (o,4)}^{drop}$)은 다음과 같이 정의된다.

(정의19) $M_{(o,1) \cup (o,4)}^{drop} = \{ M^{filtering}(O) \mid \min(\mu(x)D_A, \mu(x)D_B) \geq \alpha \}$ 이다.

예를 들어 아래 표1에서 응용 도메인 D_1, D_2, D_3 에서 비적합성 피드백을 적용한 드롭 사상은 $D_{1(0.4)} = \{o_1+o_3\}$, $D_{2(0.4)} = \{o_2+o_4\}$, $D_{3(0.4)} = \{o_3+o_4\}$ 이며, 따라서 D_1 의 드롭사상은 $M_{(o,1) \cup (o,4)}^{drop} = \{o_1+o_3\}$, D_2 의 드롭사상은 $M_{(o,1) \cup (o,4)}^{drop} = \{o_2+o_4\}$ 그리고 D_3 의 드롭사상은 $M_{(o,1) \cup (o,4)}^{drop} = \{o_3+o_4\}$ 가 된다.

3.5 퍼지 드롭 테이블

퍼지 드롭테이블은 인트라 도메인 퍼지 적합도에 따라 알파 레벨을 수행 한 후 참조 관련성 객체들을 참조 관련성 객체 클래스와 드롭 객체 클래스를 분류하기 위한 과정이다. 이때 참조 관련성 객체 클래스와 드롭 객체 클래스는 청크 블록(chunk block)으로 사상한 후 각 청크 블록들이 객체클래스들을 구성하게 된다. 청크 블록에 저장된 객체들은 필터링에 의해 동기화되며 동기화된 객체들은 퍼지 드롭 테이블에 의해 관리된다. 예를 들어 하나의 로컬 청크 블록 안에 캐싱 되는 k개의 객체가 존재하고, 객체 o_i 는 크기 s_i 를 가지고 있고, 이 청크 블록 안에는 n개의 퍼지 관련성 정보들이 존재하면 알파 레벨이 수행된 후 객체들은 퍼지 관련성에 따라 참조 관련성 객체 테이블과 드롭 테이블로 청크 블록이 재배치 되게 된다. 그림2는 알파 레벨을 수행하기 위한 청크 블록 구조이며, 이에 대한 참조 관련성 테이블과 드롭 테이블은 표1, 표2, 표3과 같다.

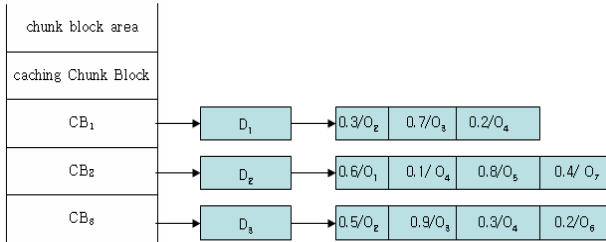


그림 2. 청크블록
Fig. 2. Chunk Block

표1. 참조 관련성 테이블

Table.1 Reference Relevance Table

D \ O	o ₁	o ₂	o ₃	o ₄
D ₁	0.3	0.7	0.2	
D ₂	0.6	0.1	0.8	0.4
D ₃	0.5	0.9	0.3	0.2

표2. $D_{1(0.1)} \cup D_{3(0.4)}$ 일 때 드롭 테이블

Table2. Drop Table with $D_{1(0.1)} \cup D_{3(0.4)}$

D \ O	o ₁	o ₂	o ₃	o ₄
D ₁	0.3		0.2	
D ₂		0.1		0.4
D ₃			0.3	0.2

표3. $D_{1(0.5)} \cup D_{3(0.9)}$ 일 때 참조 테이블

Table3. Reference Table with $D_{1(0.5)} \cup D_{3(0.9)}$

D \ O	o ₁	o ₂	o ₃	o ₄
D ₁		0.7		
D ₂	0.6		0.8	
D ₃	0.5	0.9		

청크 블록에 따라 테이블이 재배치된 후에는 퍼지 관련성에 의한 피드백 적합성을 판단하게 되며, 퍼지 관련성에 따라 0~0.5이하인 경우에는 비적합, 0.5~0.7인 경우에는 적합, 0.8이상인 경우에는 매우 적합으로 적합성 피드백을 수행한다.

4. 시스템 성능평가

이 장에서는 제안된 기법에 대해서 시스템 성능 평가를 수행한다. 시스템 성능평가의 목적은 제안된 시스템의 성능을 알아보기 위한 것이며 시뮬레이션을 위한 분산 네트워크 토폴로지, 객체 분포 및 성능 척도 등의 시뮬레이션 환경에 대해서 살펴본다.

4.1 시뮬레이션 환경

4.1.1 분산 네트워크 환경

기본적인 분산 네트워크의 환경은 유클리드 공간 모델에 기반을 두며 유클리드 공간은 D-차원 유클리드 공간에서 하이퍼큐브이다. 객체 노드들은 하이퍼큐브에 무작위로 존재하며 거리척도는 노드들 사이의 유클리드 거리 척도에 대응된다. 실험을 위한 하이퍼큐브의 에지 경로 수행 시간은 3000ms로 제한하고 알파 레벨 기법을 적용 한 후 20개의 그룹을 생성하여 시뮬레이션을 수행하였다. 그룹안의 객체들의 경우 하나의 객체에서 다른 객체들로의 전송시간은 매우 적다고 가정한다. 다른 그룹 안의 객체들의 경우에는 객체들을 전송하는 데 걸리는 시간은 그룹 들 사이의 거리에 의해 수행된다고 가정한다.

분산서비스의 신뢰성은 패킷 손실에 의해 좌우되며, 통신 경로 상에서 경로 수가 많으면 많을수록 패킷 손실률은 높아지게 된다. 그룹에서 객체들 사이의 경로 수는 객체들 사이의 거리에 비례하며, 최대 경로 수는 10으로 제한한다.

4.1.2 객체 분포

시뮬레이션을 위해서 멀티미디어 콘텐츠로서 비디오 클립과 인터넷상에서 캡처한 이미지를 사용하며, 분산 응용도메인 상에 10,000개의 객체들이 있다고 가정한다. 각 멀티미디어 객체의 비트율은 1.28Mbps로 인코드 되어 있고 각 객체 당 상연시간은 1분 이내로 제한하였으며, 크기는 5~20MB 이내로 제한하였다.

M개의 멀티미디어 객체 o_1, o_2, \dots, o_n 이 크기순으로 내림차순으로 구성되어 있다고 가정할 때 도메인의 거리 최소화를 위한 radius는 r_1 이며, $\sum_{i=1}^n r_i = 1$ 이다. 이와 같은 시뮬레이션 환경에 따라 본 논문에서는 우리의 기법과 Zipf 기법, TGD(Truncated Geometric Distribution) 기법 그리고 TPD(Truncated Pareto Distribution) 기법과의 성능 평가

를 수행하였다.

i) Zipf 분포기법

이 기법은 콘텐츠의 분포를 알아보기 위한 기법으로서 Napster[9], Gnutella[10], Freenet[11] 등에서 적용되고 있으며, 거리 $r_i = \sum_{i=1}^n i^{-(1-\alpha)} / C$ 이다. 여기서 C는 $\sum_{i=1}^n i^{-(1-\alpha)}$ 이며, α 는 skew factor 제어인수이다. 분산 멀티미디어 서비스의 경우 일반적으로 skew factor는 0.271을 적용하고 있으며, skew factor 값은 분포의 관계성을 나타낸다. skew factor가 크면 클수록 관계성이 적으며, 반대로 skew factor가 작으면 작을수록 관계성이 크게 된다. 이때 skew factor가 1이 되면 Zipf 분포는 일정하게 분포되어 있음을 의미하며, 모든 객체들은 같은 속성 및 관계성을 가지게 된다.

ii) TGD 기법

이 기법 또한 콘텐츠 분포를 알아보기 위한 기법이며 분포 계산은 $r_i = SN^i / B$,

$$B = \sum_{i=1}^n SN^i = (SN - SN^{n+1}) / (1 - SN)$$

이다. 분포 제어를 위한 임계값 SN은 0.9995로 설정하고 있다.

iii) TPD 기법

TPD 기법의 분포계산은 $r_i = i^{-\beta-1} / B_{\beta+1, N}$ 이며, 여기서

$$B_{\beta, N} = \sum_{i=1}^n i^{-\beta}$$

는 정규화 인수이고 β 는 0.01인 제어인수이다.

4.1.3 도메인 스토리지 및 신뢰도

각 도메인들은 멀티미디어 분산 서비스를 위해서 자원을 분산하게 된다. 하나의 도메인 용량은 500MB에서 2GB 사이의 정규분포를 가진다고 가정하고, 도메인 내에서의 객체는 1000개 이내로 제한하였다. 이때 신뢰도는 도메인의 적합도를 측정하는 중요한 척도로서 퍼지 적합성 척도를 위한 신뢰도는 0.5에서 1.0사이로 설정하였다.

4.1.4 성능척도

AFDM은 서비스 성능을 향상시켜, 도메인에서 관련된 객체를 사용자에게 제공하기 위한 것이다. 성능 척도를 위해 본 논문에서는 평균 접근실패율, 객체에 대한 서비스 품질, 서비스율에 대한 척도를 수행한다. 평균 접근 실패율이란 서비스 사용자가 객체를 요구한 후의 객체를 서비스하는데 걸리는 평균 접근 시간을 의미하며, 평균 접근 시간이 작을수록 전송효율이 높게 된다. 객체에 대한 서비스 품질은 클라이언트에서 서비스되는 멀티미디어 객체에 대한 품질 척도로서 객체에 대한 접근 실패율을 의미한다. 평균 접근 실패율(Access_{error})은 객체가 서비스 되지 않을 오류 ACCESS_{object}에 의해 수행되며, 평균 접근 실패율은 다음과 같다.

$$ACCESS_{error} = \sqrt{1/N \times M \sum_{x=0}^N \sum_{y=0}^M [ACCESS_{object} - ACCESS_{object}^2]}$$

서비스 품질은 접근 실패율에 의해 결정되며, 접근실패율이 적으면 객체 서비스 품질은 높아지게 된다. 그리고 서비스율은 미리 정의된 도메인의 그룹 내에서 관련된 객체들을 서비스하기 위한 척도이며 서비스율(λ)는 다음과 같다.

$$\lambda = \sum_{o_i} [r_i \times (1 - \prod_{\pi, j} (1 - p_{i,j}))]$$

여기서 p는 그룹 내에서 관련된 객체가 서비스 될 확률이다.

4.2 성능평가

이 절에서는 제안된 기법을 평가하기 위해서 평균 접근 실패율과 서비스율에 대한 실험 평 실험평가를 위해 도메인의 객체 수를 2000개에서 4000, 6000, 8,000 그리고 10000개로 2,000개 단위로 분류하여 평균 전송효율을 평가한다. 그림3에서 보듯이 우리의 기법은 객체의 수가 증가할 때 전송 효율이 크게 변하지 않지만 나머지 기법들은 평균 전송 효율이 감소하게 된다. 그리고 저장용량이 고정된 경우에 새로운 객체를 추가하게 되면 나머지 기법들은 새로운 객체에 대한 비교가 적절하게 제공되지 않아 검색효율이 감소되는데 비해서 우리의 기법은 알파 레벨에 의한 드롭 사상이 수행되어 검색 효율에 영향을 크게 미치지 않게 된다.

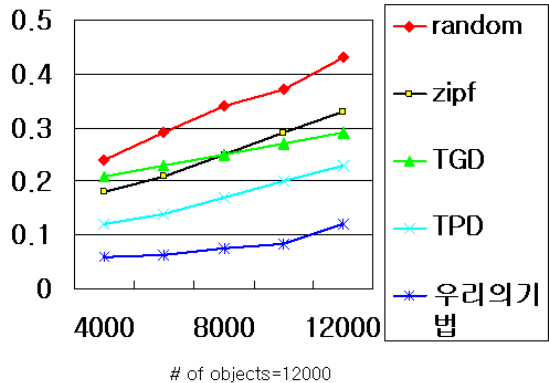


그림 3. 평균 접근 실패율
Fig. 3. Average access fault rate

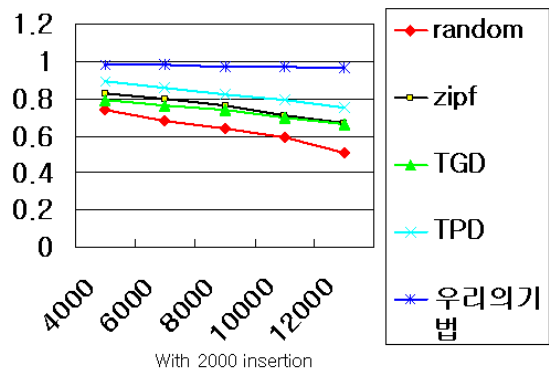


그림 4. 2000의 객체를 추가하였을 때의 평균 검색율
Fig. 4. Average access fault rate with 2000 insertion

그림4는 도메인에서 새로운 객체 2000개를 추가하였을 때의 평균 검색율이다. 그림에서 보듯이 응용도메인에서 추출된 객체를 드롭사상에 따라 분류된 그룹에는 검색율에 큰 변화를 주지 않지만 그렇지 않는 경우에는 검색효율이 더 떨어짐을 알 수 있게 된다. 서비스율은 객체의 분포를 평가하기 위한 모델로서 제어 인수를 이용하여 성능을 측정하게 된다. 그림4에서 TGD의 성능 평가를 위한 임계값 SN은 0.9995로 설정하였으며, TPD의 임계값은 제어 인수는 0.01

로 설정하였다. 그리고 우리의 기법은 적합과 매우 적합한 적합성 피드백을 적용하여 평가를 수행하였다.

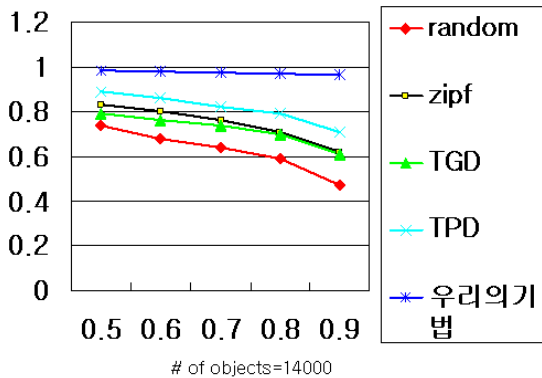


그림 5. 평균 서비스율
Fig. 5. Average service rate

그림5에서 보듯이 도메인 객체 수가 증가될 때 Random 기법 및 다른 기법은 선형으로 감소됨을 알 수 있으며, 우리의 기법은 객체의 수가 증가될 때 평균 서비스율이 증가됨을 알 수 있다. 따라서 우리의 기법은 도메인의 객체 수가 12000에서 14000개로 증가될 때 평균 서비스율이 다른 기법에 비해서 성능이 향상됨을 알 수 있다.

5. 결 론

최근에 분산 모바일 및 무선 서비스가 확대됨에 따라 분산 환경에서 사용자 위주의 서비스를 제공하기 위한 여러 기법들이 제안되고 있다[15]. 본 논문에서는 이러한 요구에 따라 사용자의 관심도 및 서비스율, 검색시간을 개선하고, 수작업을 통한 기법에 비해 서비스 성능을 향상시킬 수 있는 신뢰성 있는 AFDM을 제안하였다. 제안된 시스템은 인터넷 도메인 적합도와 인프라 도메인 적합도에 의해 신뢰성이 보장되도록 하였으며, 인터넷 도메인 적합도를 위해 KRF에 의한 거리 최소화 기법을 적용하였다. 그리고 인터넷 도메인 적합도는 알파 레벨에 의한 퍼지 적합도를 적용하여 도메인 객체에 대한 적응성을 보장하도록 하였다. 또한 질의 확장을 통하여 객체 추론이 적응 가능하도록 하였으며, 사상스케줄러를 통하여 참조 관련성이 적은 객체들을 드롭하였다. 사상스케줄러에 의해 사상된 객체들은 퍼지 드롭 테이블에 의해 참조관련성 객체 클래스와 드롭 객체 클래스로 관리되도록 하였다. 시뮬레이션 결과 AFDM에 의해서 관리된 객체들은 성능 향상에 큰 변화를 가져옴을 알 수 있으나 그렇지 않는 경우에는 검색 성능이 떨어짐을 알 수 있었다. 따라서 제안된 AFDM 기법이 검색효율성 및 서비스 효율성 측면에서 성능이 향상됨을 알 수 있었다.

참 고 문 헌

[1] Songqing Chen, Bo Shen, Susie Wee, and Xiaodong Zhang, "Segment-Based Streaming Media Proxy: Modeling and Optimization," *IEEE*

TRANSACTION ON MULTIMEDIA, VOL. 8, NO. 2, pp. 243-256, 2006.
 [2] Peng Zhu, Wenjun Zeng, and Chunwen Li, "Joint Design of Source Rate Control and QoS-Aware Congestion Control for Video Streaming Over the Internet," *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 9, NO. 2, pp.366-376, 2007.
 [3] Chi-Feng Kao and Chung-Nan Lee, "Aggregate Profit-Based Caching Replacement Algorithms for Streaming Media Transcoding Proxy Systems," *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 9, NO. 2, pp.221-230, 2007.
 [4] J.Kangasharju, FHartanto, M. Reisslein, and K. W. Ross, "Distributing layered encoded video through caches," *IEEE Trans. Comput.*, vol. 51, pp.622-636, 2002.
 [5] Zhe Xiang, Qian Zhang, Wenwu Zhu, Zhensheng, and Yu-Qin Zhang, "Peer-to-Peer Based Multimedia Distribution Services," *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 6, NO. 2, pp.343-354, 2004.
 [6] D. A. Tran, K. A. Hua, T. T. Do, " A peer-to-peer architecture for media streaming," *IEEE J.Select. Areas Comm., Special Issue on Recent Advances in Service Overlay Networks*, vol. 22, pp.121-133, 2004.
 [7] S.Chen, B.Shen, S.We, and X. Zhang, " Adaptive and Lazy Segmentation based Proxy Caching for Streaming Media Delivery, " *in Proc. ACM NOSSDAV*, Monterey, CA, pp.429-441, 2003.
 [8] [online]. Available: <http://www.alkami.com/>
 [9] [online] Available: <http://www.napster.com/>
 [10] [online] Available: <http://www.gnutella.com/>
 [11] [online] Available: <http://freenet.sourceforge.com/>
 [12] [online] Available: <http://www.centerspan.com/>
 [13] Z. Zang, Y. Wang, D. Du, and D. Su, "Video staging: A Proxy Server based Approach to end-to end Video Delivery over Wide Area Networks," *IEEE Trans. Networking*, vol. 8, pp.429-442, 2000.
 [14] 유동선, 이교원, *기초퍼지이론*, 교우사, 2006.
 [15] Wei Wang and Chinya V. Ravishhankar, "Adaptive Broadcasting for Similarity Queries in Wireless Content Delivery Systems," *IEEE TRANSACTIONS ON Knowledge and Data Engineering*, VOL. 20, NO. 4, pp.504-518, 2008.

저 자 소 개



정택원(Taegwon Jeong)

1981년 2월 : 서울대학교 대학원 전기공학과 (공학석사)

1991년 : Univ. of Florida Dept of EE (Ph.D)

1983년~1998년 : ETRI 책임연구원

1998년~현재 : 전북대학교 응용시스템 공학부 교수

관심분야 : 정보통신, 이동통신

E-mail : ttwjeong@yahoo.co.kr



이종득(Chong Deuk Lee)

1989년 : 전북대학교 대학원 전산통계학과 (이학석사)

1998년 : 전북대학교 대학원 전산통계학과 (이학박사)

1992년~2002년 : 서남대학교 컴퓨터정보통신학과 교수

2002년~현재 : 전북대학교 응용시스템공학부 교수

관심분야 : Mobile Ad-hoc Networks, Sensor Networks, 멀티미디어 통신, Mobile 성능평가

E-mail : cdlee1008@chonbuk.ac.kr