

논문 2007-44CI-3-2

상황 인식 환경에서 온톨로지를 이용한 프로액티브 검색 기법

(Proactive Retrieval Method using Ontology
in Context-aware Environment)

김 성 립*, 권 준 희**

(Sungrim Kim and Joonhee Kwon)

요 약

상황 인식 환경에서는 물리적인 환경, 상황 등을 시스템이 인식하고 이를 기반으로 사용자와의 상호작용을 지원하는 상황 인식 기술이 중요한 요소로 자리잡고 있다. 온톨로지 기반의 상황 정보 모델은 상황정보의 공유와 재사용의 이점을 제공하기 때문에 최근 널리 사용되고 있다. 본 논문에서는 상황 인식 환경에서 온톨로지를 이용한 새로운 프로액티브 검색 기법을 제안한다. 제안된 기법은 온톨로지의 계층적 개념 트리를 이용하여 사용자의 상황에 맞는 정보의 개념 수준을 결정함으로써 보다 효율적인 검색이 가능하다. 이를 위해 제안된 기법을 설명하고, 실험을 통해 기존 기법보다 제안 기법이 우수함을 보인다.

Abstract

The context-aware environment focuses on recognizing the context and physical entities. For this reason, there has been an increasement in research of context-aware computing environment. Ontology-based context models are widely used in ubiquitous environment because of context sharing and reusing. In this paper, we propose a proactive retrieval method using ontology in context-aware environment. The method use a concept level of hierarchical concept tree in ontology for more efficient retrieval. This paper describes the proactive retrieval method and ontology model. Several experiments are performed and the results verify that the proposed method's efficiency is better than other existing methods.

Keywords : context-aware, proactive, ontology, retrieval

I. 서 론

유비쿼터스 환경에서는 기존 컴퓨팅 환경과는 달리 상황(context)을 시스템이 인식하고 이를 기반으로 사용자와 상호 작용하는 상황 인식 기술이 필수적인 요소로 인식되고 있다. 또한 사용자의 명시적인 요구에 따라 제공되는 정보보다는 각 사용자의 상황에 따라 프로

액티브하게 정보를 제공하는 기술이 요구된다^[1].

상황 인식 환경에서 언제 어디서나 상황에 따른 정보가 올바르게 제공되기 위해서는 정보에 의미가 부여되어야 하며 정보들이 공유되어야 한다. 이러한 필요에 따라 상황 인식 환경에서 온톨로지를 활용하는 연구가 최근 증가하고 있다. 그루버는 온톨로지를 공유된 개념에 대해 정형화되고 명시적으로 정의된 명세서로 정의하고 있다^[8]. 온톨로지는 정보의 의미를 부여하고, 정보들 간의 관계를 설정해 줌으로써 정보를 효과적으로 표현하고 공유하는 것을 그 목적으로 한다^[2,7].

그러나 상황 인식 환경에서 온톨로지를 이용한 검색 기법에 대한 기존 연구들에서는 검색되는 정보의 개념

* 정회원, 서일대학 인터넷정보과
(Dept. of Internet Info. Seoil College)

** 정회원, 경기대학교 정보과학부 전자계산학전공
(Dept. of Computer Science, Kyonggi Univ.)

※ 본 논문은 2006년 서일대학 학술연구비의 지원에 의해 연구되었음.

접수일자: 2007년1월25일, 수정완료일: 2007년5월4일

수준이 상황별로 다르다는 점을 간과하여 왔다. 이에 따라 사용자에게 불필요한 정보가 제공되어 검색 시간의 증가와 함께 원하는 정보를 찾는 데 상당한 노력과 시간을 소요하게 될 수도 있다. 사용자의 상황은 계속적으로 변화하며 각 상황에 따라 사용자가 제공받고자 하는 정보의 개념 수준은 모두 다르다. 이를 위해서는 각 상황에 필요한 수준의 정보만을 제공함으로써 계속적으로 변화하는 상황에도 유용한 정보만을 신속하게 검색할 수 있어야 한다.

본 논문에서는 온톨로지의 계층적 개념 트리를 이용하여 사용자의 상황에 맞는 정보의 수준을 결정하여 프로액티브하게 정보를 검색할 수 있는 기법을 제안한다. 제안 기법에서는 상황을 온톨로지를 기반으로 모델링하여 상황에 대한 공통의 이해를 통해 다른 상황 인식 환경에서도 정보를 검색할 수 있다. 그리고 사용자의 상황과 관심 정도에 따라 개념 수준을 결정하여 사용자들에게 보다 유용한 정보가 검색될 수 있도록 한다.

본 논문의 구성은 다음과 같다. II장에서는 관련 연구를 살펴보고, III장은 온톨로지를 이용한 프로액티브 검색 기법을 설명한다. IV장에서는 제안 기법을 실험 평가하고 마지막으로 V장에서 결론을 맺는다.

II. 관련 연구

최근 유비쿼터스 환경의 도래와 함께 데스크톱 컴퓨터만으로는 부족했던 유용한 서비스를 사용자에게 제공할 수 있는 환경이 구축되고 있다. 그러나 이와 관련한 대부분의 연구는 사용자의 상황을 고려하지 않아 상황에 맞게 정보를 검색하여 서비스를 제공할 수 없다는 한계가 있다.

유비쿼터스 환경에서는 끊임없이 변화하는 사용자의 상황을 고려하는 상황 인식 검색 서비스가 매우 중요한 연구 중 하나로 대두되고 있다. 이러한 연구들 중 가장 최근의 연구 경향 중 하나로는 시맨틱 웹을 활용하여 상황에 따라 보다 유용한 정보를 검색하고자 하는 연구를 들 수 있다. 이러한 연구들 중 대표적인 몇 가지를 살펴보면 다음과 같다.

COMPASS^[3]는 사용자의 흥미와 현재 위치에 기반하여 사용자에게 필요한 서비스를 추천하는 상황 인식 모바일 관광 서비스이다. 온톨로지를 사용하여 상황에 필요한 정보를 제공하고, 이를 통해 지식 공유와 추론이 가능해져 각 상황에 유용한 정보를 제공할 수 있다.

CRUMPET^[4]은 온톨로지와 추론을 사용하여 사용자

친화적이고 개인화가 적용된 관광 추천 서비스를 목적으로 한다. 따라서 온톨로지를 사용함으로써 사용자에게 의미적으로 가장 가까운 정보를 추천할 수 있다.

MyCampus^[5]와 Cool Agent Recommendation Service^[6] 역시 사용자의 상황을 온톨로지로 모델링하고 규칙을 정의하여 추론된 결과를 권유하는 추천 시스템이다. 이 연구들에서는 온톨로지를 이용하여 시스템 혹은 에이전트간에 정보를 교환하고 추론이 가능하도록 함으로써 보다 효과적으로 정보를 제공할 수 있다.

그러나 이러한 기존 연구들은 사용자의 상황과 관심 정도에 따라 제공되어야 할 정보의 개념 수준이 서로 다르다는 부분을 고려하지 않고 있다. 이는 사용자의 상황에 불필요한 정보를 제공함으로써 정보 제공의 신속성과 유용성 측면에서 문제를 야기한다.

III. 온톨로지를 이용한 프로액티브 검색 기법

1. 프로액티브 검색 기법 개요

그림 1은 상황 인식 환경에서 온톨로지를 이용한 프로액티브 검색 기법을 보여준다. 본 논문에서는 상황에 기반해서 사용자에게 정보를 효율적으로 서비스하기 위해 클라이언트 에이전트와 서버 에이전트를 사용한다.

클라이언트 에이전트는 사용자의 상황 정보를 수신하여 상황별 관심도를 측정 후 필요한 정보를 서버 에이전트에 요청하여 개념 수준별로 검색한 정보를 제공한다. 서버 에이전트는 수집된 상황 정보 데이터베이스를 이용하여 선호도를 측정한다. 측정된 선호도를 사용해 온톨로지의 계층적 개념 트리를 사용한 정보 개념

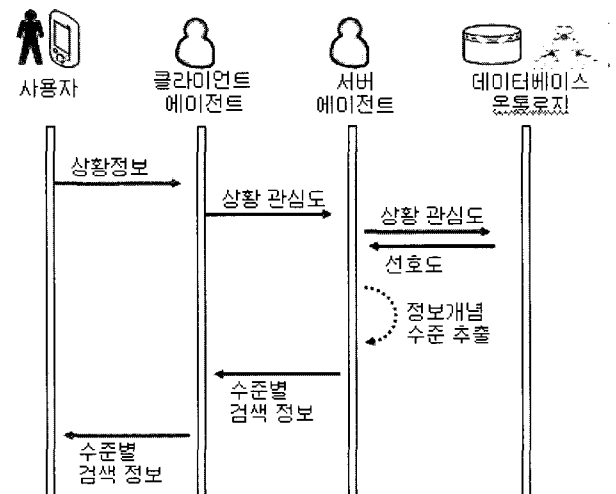


그림 1. 프로액티브 검색 기법 흐름도
Fig. 1. Flow diagram of proactive retrieval method.

수준의 결정을 내릴 수 있으며 측정된 상황 관심도에 따라 검색할 정보의 양을 결정하게 된다.

클라이언트 에이전트는 변경되는 상황에 따라 정보를 빠르게 전달할 수 있도록 프리페칭과 캐싱 기법을 사용한다. 즉, 현재 위치 상황에 대한 관심도가 높은 경우 해당 위치에 포함된 다른 위치들은 가까운 미래에 도달할 위치로 예상이 가능하므로, 이에 대한 정보를 미리 캐시에 프리페칭한다. 또한, 과거에 검색된 정보는 미래에도 검색될 가능성이 높으므로 캐시에 저장한다. 이를 통해, 해당 위치 상황에 도달했을 때 이에 대한 정보가 캐시에 있는 경우 서버에게 요청하지 않고 정보를 추천할 수 있어 보다 신속한 검색이 가능하다.

2. 계층적 개념 트리

본 논문에서는 그림 2와 같이 상황 인식 환경에서 프로액티브하게 정보를 검색하기 위해 온톨로지에 기반을 두어 전자상거래 도메인을 디자인하였다. Person, StoreProduct, StoreLocation, History 4가지 도메인 타입의 OWL 파일로 이루어져 있으며 'subclass' 관계는 실선으로, 'property' 관계는 점선으로 표시하였다.

Person 온톨로지는 서비스를 사용하는 사용자를 나타낸 온톨로지로서 사용자 각각에 대한 정보를 담아 사용자에 대한 식별자 역할을 한다. StoreProduct 온톨로지는 상품을 개념적으로 동질한 것으로 분류시킨 것이며 Product(상품) 클래스의 하위 클래스이다. StoreLocation 온톨로지는 매장의 위치와 매장간의 의미적 집합 관계를 나타낼 수 있도록 구성되었다. History 온톨로지는 사용자가 매장에서 과거에 구매했던 구매 기록을 표현하기 위한 온톨로지로서 다른 모든 온톨로지 도메인을 참조하여 구매 기록을 저장하여 과거 구매 선호도를 계산할 수 있게 한다.

본 논문에서는 언제 어디서나 상황의 변화에 관계없

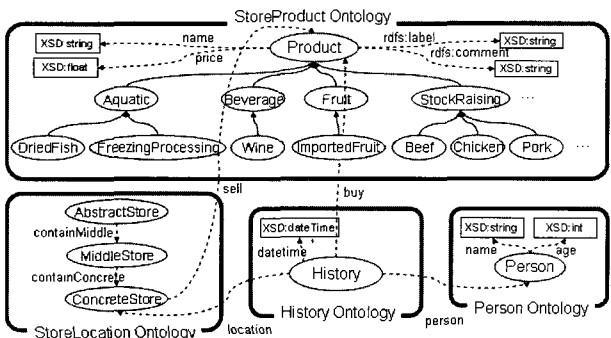


그림 2. 전자상거래 온톨로지
Fig. 2. E-commerce Ontology.

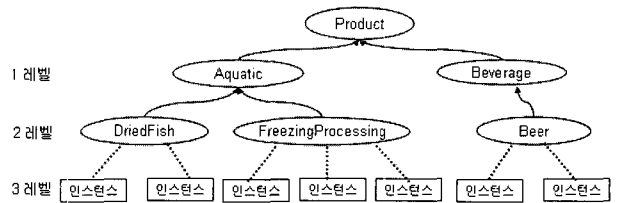


그림 3. 상품 온톨로지
Fig. 3. Product Ontology.

이 정보를 제공하기 위해서 온톨로지를 사용하였다. 온톨로지를 사용함으로써 공통의 지식을 공유하게 되어 사용자가 처음 방문하는 곳이라도 온톨로지에 기반을 두어 지식을 공유하기 때문에 사용자에게 유용한 정보를 제공할 수 있게 된다.

또한 기존의 연구와 차별화하여 온톨로지의 계층적 개념 트리 특성을 활용하였다. 이를 통해, 각 매장 위치마다 사용자에게 추천할 정보의 개념 수준을 결정하고 사용자에게 보여줄 정보의 양을 결정함으로써 보다 유용한 정보가 제공될 수 있다.

그림 3은 그림 2에 나타난 상품 온톨로지를 보다 상세화해서 나타내고 있으며, 이를 통해 계층적 개념 트리에 따라 상품 정보의 개념 수준이 달라지는 것을 알 수 있다. 상품 정보는 온톨로지 개념 트리를 따라 상위 수준으로 올라갈수록 일반적인 의미의 개념 정보가 되며 정보의 양이 감소하고, 하위 수준으로 내려갈수록 구체적 수준의 상품 정보가 되며 정보의 양이 증가하는 특징이 있다.

본 논문에서 제안하는 검색 기법은 크게 3가지 흐름으로 구성된다. 첫째, 사용자의 구매 기록을 바탕으로 사용자가 선호하는 상품을 추출한다. 상품 온톨로지 개념 트리의 상품 선호도는 인스턴스별 상품 선호도를 구한 후 1레벨씩 상위 수준의 개념으로 올라가면서 클래스별 상품 선호도를 구하게 된다. 이러한 절차에 따라 우선 계산되는 한 인스턴스 i 에 대한 상품 선호도 $ProductPreference_i$ 는 식 1과 같다.

$$ProductPreference_i = \frac{\text{상품 } i \text{의 구매횟수}}{\text{전체 상품 구매횟수}} \quad (1)$$

이러한 인스턴스별 선호도를 바탕으로 상품 온톨로지 개념 트리의 클래스 노드 i 에 대한 상품 선호도를 식 2와 같이 계산한다.

$$ProductPreference_i = \frac{\text{클래스 } i \text{를 부모노드로 하는 자식노드들의 상품 선호도의 합}}{\text{클래스 } i \text{의 높이 트리 최대 높이}} \quad (2)$$

식 1과 식 2를 통해 각 상품 개념 트리 노드별 선호도를 모두 구하게 되면, 이 중 상품 선호도의 최소 임계값보다 큰 상품만을 추출한다. 이러한 과정을 통해 사용자가 선호하는 상품을 추출하게 된다.

둘째, 사용자에게 제공할 정보의 개념 수준을 결정한다. 본 논문에서 제안한 기법에서는 첫 번째 단계를 통해 추출된 인스턴스 레벨의 선호 상품을 각 위치 상황별로 단순히 제공하지 않는다. 이는 사용자가 한번에 살펴볼 수 있는 정보의 양은 제한되어 있으며 이를 초과하게 되면 상세 정보가 아닌 추상적 정보를 통해 정보의 양을 줄여 정보를 얻게 된다는 기초적인 개념에 근거한다.

개념 수준을 결정하기 위해 두 번째 흐름에서는 매장 위치별로 보유하고 있는 사용자 선호 상품 개수와 사용자가 보유한 디바이스 장치에 제공할 수 있는 정보의 최대 개수를 주요 요소로 사용한다. 즉, 사용자의 현재 위치 상황에서 선호하는 상품 정보의 양이 사용자 디바이스에서 보여줄 수 있는 정보의 양을 초과하게 되면 개념 트리의 개념 레벨을 상향시켜 보다 일반화된 개념 정보를 제공하게 된다. 이를 통해 정보의 양을 줄이면서도 필요한 모든 정보를 얻을 수 있게 된다.

셋째, 사용자의 위치 상황별 관심도에 따라 제공할 정보의 양을 조절하여 이를 사용자에게 제공한다. 두 번째 단계를 통해 정보의 개념 수준을 결정한 후에는 사용자의 현재 위치 상황에 대한 관심 정도에 따라 추천할 정보의 양을 결정한다. 현재 위치 상황에 대한 관심 정도는 과거 위치 상황에 대한 선호도와 현재 위치 상황에 대한 관심 정도를 사용하여 계산된다. 이러한 요소를 이용한 현재 상황 j 에 대한 관심도 $AttentionPreference_j$ 는 과거 상황 j 에 대한 선호도 $LocationPreference_j$ 와 현재 상황 j 에 대한 관심 정도 $BehaviorPreference_j$ 를 이용하여 식 3과 같이 계산한다.

이 때 α 값은 $AttentionPreference_j$ 와 $BehaviorPreference_j$ 의 비율을 조절할 수 있는 임계값이다.

$$AttentionPreference_j = (1-\alpha) \times LocationPreference_j + \alpha \times BehaviorPreference_j \quad (0 \leq \alpha \leq 1) \quad (3)$$

식 (3)에서 과거 상황 j 에 대한 선호도 $LocationPreference_j$ 는 식 (4)에 의해 계산된다.

$$LocationPreference_j = \frac{j\text{위치의 총 구매횟수}}{\text{모든 위치의 총 구매횟수}} \quad (4)$$

현재 상황 j 에 대한 관심 정도 $BehaviorPreference_j$ 는 예를 들면, 사용자의 현재 위치에서의 속도 등의 요소에 의해 알 수 있다. 즉, 사용자가 서 있는 경우는 현재 상황에 관심이 높다고 판단하여 관심 정도를 1이라고 하며 걷는 경우를 0.5, 뛰는 경우를 0이라고 설정할 수 있다.

현재 상황에 대한 관심도를 구하면 사용자 디바이스에서 제공할 수 있는 정보의 최대 양을 기준으로 현재 상황에 대한 관심 정도에 맞는 정보의 양만큼을 제공하게 된다. 세 번째 단계를 통해 정보의 양이 결정되면 최종적으로 사용자에게 정보를 제공하게 된다.

IV. 실험

1. 실험 환경

제안 기법의 성능을 평가하기 위해 제안 기법과 기존 기법을 구현하여 실험하였다. 기존 기법은 제안 기법과는 다르게 프리패칭, 정보의 개념 수준 변경, 그리고 사용자의 위치 상황별 관심도에 따라 제공할 정보의 양을 조절하는 절차없이 구현하였다.

구현 환경은 다음과 같다. 서버측 환경은 데이터베이스 MySQL 5.0.18, 온톨로지 모델링 도구 Protege 3.2 beta, 시맨틱 웹 라이브러리 Jena 2.4, 자바 가상머신 1.4.2, SOAP 엔진 Axis 1.4, 온톨로지를 위한 쿼리 언어 SPARQL을 사용하였다. 클라이언트는 Windows Mobile 2003 SE 에뮬레이터에서 Visual C# .Net으로 구현하였다.

실험에 사용되는 온톨로지는 3장에서 사용한 온톨로지를 그대로 사용하였다. 이 중 상품 온톨로지서 보유하고 있는 상품 인스턴스의 총 개수는 420개로 하였다. 매장 데이터는 온톨로지를 공유하는 10개의 매장을 대상으로 하여, 각 매장별로 420개의 상품을 10%씩 증가하면서 10~100%에 걸쳐 무작위로 보유하도록 데이터를 구성하였다. 또한 매장별 보유 상품 중 사용자가 선호하는 선호도를 20%, 30%, 40%, 50%로 나누어 실험을 수행했는데 이 때 선호하는 상품은 무작위로 하여 데이터를 구성하였다. 이 때, 서로 다른 100명의 사용자를 가상하여 데이터를 생성하였다.

사용자의 구매 비율은 현재 매장에서 판매하는 상품 중 사용자가 선호하는 상품의 20%개수만큼을 비선호 상품으로 구매하고, 나머지 80%를 선호상품으로 구매한다고 가정하였다. 또한 선호 상품은 상품 선호도 상위 20%에서 80%를 구매하고, 하위 80%에서 나머지를

무작위로 구매한다고 가정하였다. 사용자가 방문하는 매장에서 상품을 구매하는 경우는 관심선도 값을 대매장인 경우는 0.5, 중매장과 소매장인 경우는 1.0이라고 가정하였다. 구매하지 않는 경우에 대매장과 중매장 그리고 소매장의 관심선도 값은 모두 0.2이라고 가정하였다. 마지막으로 사용자는 같은 상품을 두번 구매하는 일은 없으며 모든 소매장을 단 한 번씩만 방문하고 중매장과 대매장은 소매장을 방문하기 위해 거쳐야 하는 최소한의 방문만을 한다고 가정하였다.

2. 실험 결과

효율성(eficiency)을 실험함으로써 기존 기법(System1)과 제안 기법(System2)의 성능을 평가하였다. 이를 위해 사용자의 상황이 변한 후 사용자에게 필요한 정보를 제공하는데 걸리는 평균 응답 시간을 측정하였다. 효율성 측정을 위해서는 캐시 크기를 변경하여 실험하였다. 이에 대한 실험 결과는 그림 4와 같다.

실험 결과를 통해 캐시 크기가 클수록, 상품 선호도가 증가할수록 System2의 성능이 더 좋아짐을 알 수 있는데 그 이유는 다음과 같이 분석된다. 첫째, 캐시 크기가 증가할수록 System2는 프리패칭할 수 있는 정보의 양이 증가하여 서버와의 정보 전송량이 줄어들게 된다. 둘째, 상품 선호도가 증가할수록 사용자에게 제공할

정보의 양이 늘어나게 되는데 이 때 해당 정보 모두를 검색하는 System1에 비해, System2는 이를 개념 수준의 변경과 각 위치 상황별 관심도에 따라 해당 정보의 양이 줄어들기 때문이다.

V. 결론

오늘날 상황 인식 환경에서 사용자에게 유용한 정보를 프로액티브하게 검색하여 제공할 수 있는 서비스와 상황에 따라 제공되는 정보에 의미를 부여함으로써 보다 유용한 정보를 제공하고자 온톨로지를 이용하는 연구가 증가하고 있다.

그러나, 기존의 연구들은 사용자의 상황과 관심 정도에 따라 제공되어야 할 정보의 개념 수준이 서로 다르다는 부분을 고려하지 않는다는 문제점이 있다. 이는 사용자의 상황에 불필요한 정보를 제공함으로써 정보 제공의 신속성과 유용성 측면에서 문제를 야기한다. 이를 위해 본 논문에서는 상황 인식 환경에서 온톨로지를 이용한 새로운 프로액티브 검색 기법을 제안하였다. 이는 사용자가 한번에 살펴볼 수 있는 정보의 양은 제한되어 있으며 이를 초과하게 되면 상세 정보가 아닌 추상적 정보를 통해 정보의 양을 줄여 정보를 얻게 된다는 기초적인 개념에 근거한다.

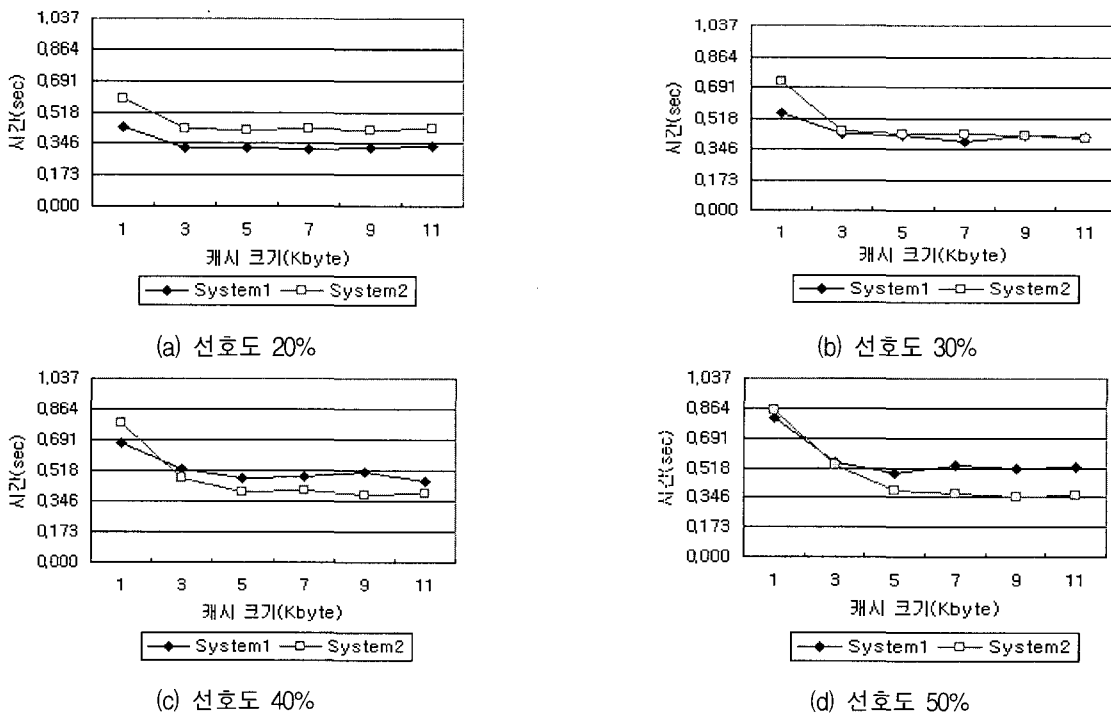


그림 4. 평균 응답 시간
Fig. 4. Average Response Time.

본 논문의 공헌도는 다음과 같다. 첫째, 상황과 제공 정보를 전자상거래 도메인을 대상으로 온톨로지 모델링 하였다. 이를 통해 상황과 제공 정보의 공유가 가능하여 한 번도 방문하지 않은 위치 상황에서도 정보 제공이 가능하다. 둘째, 사용자의 상황별 관심도에 따라 정보의 양을 조절함으로써 불필요한 정보를 제공하지 않는다.

참 고 문 헌

[1] A. Dey, D. Abowd and D. Salber, "Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications," Anchor article of a special issue on context-aware computing in the Human-Computer Interaction Journal, Vol. 16 No. 2-4, pp.97-166, 2001.

[2] Tim Berners-Lee, James Handler and Ora Lassila, "The Semantic Web," Scientific American, May 2001.

[3] Mark van Setten and Stanislav Pokraev, "Context-Aware Recommendations in the Mobile Tourist Application COMPASS", LNCS 3137, pp. 235-244, 2004.

[4] S. Poslad, H. Laamanen, R. Malaka, A. Nick, P. Buckle, and A. Zipf. "CRUMPET: Creation of User-friendly Mobile Services Personalised for Tourism". In Proceedings of Second International Conference on 3G Mobile Communication Technologies, pp.28-32, March 2001.

[5] Norman Sadeh, "A Semantic Web Environment for Context-Aware Mobile Services". In The Proceedings of Wireless World Research Forum Conference, September 2001.

[6] Harry Chen, Sovrin Tolia, Craig Sayers, Tim Finin, and Anupam Joshi. "Creating Context-Aware Software Agents". First International Workshop on Radical Agent Concepts, WRAC 2002, pp.186-200, McLean, VA, USA, January 16-18, 2002.

[7] Latifur Khan, Dennis McLeod, Eduard H. Hovy: "Retrieval effectiveness of an ontology-based model for information selection," VLDB Journal. Vol. 13, No. 1, pp.71-85, 2004.

[8] Gruber, T. "A Translation Approach to Portable Ontology Specifications," Knowledge Acquisition Vol. 5, No. 2, pp.199 - 220, 1993.

저 자 소 개



김 성 림(정회원)
 1994년 숙명여자대학교 전산학과 학사 졸업.
 1997년 숙명여자대학교 대학원 전산학과 석사 졸업.
 2002년 숙명여자대학교 대학원 전산학과 박사 졸업.

2004년 3월~현재 서일대학 인터넷정보과 교수.
 <주관심분야 : 유비쿼터스 컴퓨팅, 웹 데이터베이스, XML, 멀티미디어 의료 데이터베이스>



권 준 희(정회원)
 1992년 숙명여자대학교 전산학과 학사 졸업.
 1994년 숙명여자대학교 대학원 전산학과 석사 졸업.
 2002년 숙명여자대학교 대학원 전산학과 박사 졸업.

2003년 3월~현재 경기대학교 정보과학부 전자계산학전공 교수.
 <주관심분야 : 유비쿼터스 컴퓨팅 및 LBS, 공간 데이터베이스 및 GIS, 워크플로우>