

사용자 프로파일을 이용한 전자 상거래 검색 시스템 설계

김두현⁰ 홍도석 김용성
전북대학교 컴퓨터학과

{kdh, capcom}@mail.chonbuk.ac.kr, yskim@moak.chonbuk.ac.kr

Design of Electronic Commerce Search System Using User-Profile

Doo-Hyun Kim⁰ Do-Seok Hong Yong-Sung Kim
Dept. of Computer Science, Chonbuk National University

요 약

본 논문에서는 사용자의 요구사항을 만족하는 최적의 상품을 검색하는 전자 상거래 검색 시스템을 설계하고자 한다. 우리는 전자 상거래의 물품 구입에 대한 사용자 설문 조사를 통하여 소비자에게 영향을 미치는 요소들을 추출한다. 이러한 요소들 중에서 검색 엔진이 비교할 수 있는 요소들을 소비자 프로파일(Profile) 요소로 결정하고, 인터넷의 많은 상품 중에서 퍼지(Fuzzy) 이론을 적용하여 사용자의 프로파일에 적합한 상품 검색 알고리즘을 제안해 소비자가 상품 검색을 쉽게 할 수 있는 검색 시스템을 설계한다.

1. 서 론

컴퓨터의 대중화와 통신 기술의 발달로 인해, 우리는 인터넷에서 많은 정보를 얻을 수 있게 되었다. 인터넷과 웹의 폭넓은 사용은 우리의 생활에 많은 영향을 미치고 있다. 그 중에 하나가 전자 상거래이다. 전자 상거래는 전 세계의 사이버 공간에서 상품뿐만 아니라 정보, 콘텐츠 등을 판매한다.

전자 상거래의 활성화를 위해 물건을 쉽게 관리하기 위한 카테고리 분류, 결재를 위한 전자 화폐 운영, 사용자의 프라이버시를 위한 보안성 등에 관해서는 많은 연구가 있었다. 이러한 전자 상거래는 시간의 단축과 편리성 때문에 전자 상거래 사이트와 상품이 증가하고 있다. 그러나 많은 양의 상품은 검색을 더욱 더 복잡하게 만든다. 따라서, 상품 검색을 쉽게 하기 위해 사용자에게 새로운 상품 검색 메커니즘의 필요성이 증가하고 있다.

여러 검색 도구들을 사용하여 효과적인 검색 결과를 제공하기 위한 많은 연구들이 있었다. 그러나 이러한 연구는 사용자의 정보 검색 시간이나 관심도를 고려하지는 못해 사용자는 상품을 찾기 위해서 많은 시간을 낭비하고 피드백 과정을 거쳐야만 한다[1, 3].

일반적으로 사용자 프로파일은 사용자의 특성을 나타내는 항목들의 집합으로 구성된다. 이러한 프로파일은 사용자의 관심도를 표현하고, 이것을 이용해서 사용자에게 적합한 검색 결과를 제공할 수 있다. 또한, 사용자의 관심을 고려한 검색은 검색의 효율을 향상시킬 수 있다.

따라서 본 논문에서는 퍼지 논리를 기반으로 사용자의 관심도에 따른 전자 상거래의 상품을 검색할 수 있는 알고리즘을 제안하고, 사용자가 원하는 수준의 제품을 검색할 수 있는 전자 상거래 검색 시스템을 설계한다.

논문의 구성으로, 2장은 검색 방법에 대한 관련 연구를 소개하고, 3장에서는 사용자 프로파일을 고려한 물건 검색 알고리즘을 제안한다. 그리고 4장에서는 시스템을 설계하고, 5장에서는 결론과 향후 연구과제를 제시한다.

2. 관련 연구

전자 상거래의 상품 검색 방법에는 키워드 검색과 필드별 검색, 상품의 분류 체계에 기반으로 하는 계층적 검색 등이 있다. 이러한 방법들은 많은 연구가 이루어져 왔고 그 속도나 정확도에 있어서 큰 발전이 있었다.[2] 하지만 대상이 되는 정보가 기하급수적으로 늘어나고, 상품에 대한 정확한 분류가 어려워 소비자의 관심도에 따른 내용을 효율적으로 검색하는 정보검색에는 적합하지 못하다.[4]

또한, 최근에 많이 연구되고 있는 질의어 확장에 대한 연구는 개인, 그룹의 프로파일을 이용하는 기법과 사용자 피드백에 의한 기법이 있다.[7] 사용자 프로파일을 사용하는 경우는 사용자가 질의를 입력하면, 사용자의 관심 분야에 관련된 키워드들로 작성된 프로파일이나 시소러스를 참조하여 질의어를 확장하게 된다[3]. 위와 같은 기법들은 프로파일이나 시소러스 구축에 있어서 많은 시간과 노력이 필요하며 동적으로 이들을 유지보수하기가 무척 어려운 단점이 있다.

따라서, 본 논문에서는 사용자가 전자 상거래에서 제품을 선택하는데 영향을 미치는 요소들은 설문 조사를 통해서 추출하고, 추출된 항목 중 검색 엔진이 비교할 수 있는 것들을 선별해서 사용자의 프로파일 요소로 결정한다. 이러한 사용자의 프로파일을 이용하여 인터넷에 존재하는 많은 데이터 중 사용자의 요구 사항에 적합한 데이터를 퍼지 논리를 기반으로 선택하는 알고리즘을 제안하고, 전자 상거래 검색 시스템을 설계한다.

3. 사용자 프로파일에 따른 검색 알고리즘

본 논문에서 제안하는 사용자 프로파일에 따른 전자 상거래 검색 시스템의 구조는 그림 1과 같다.

전처리 단계로 사용자의 프로파일 요소는 메일에 의한 설문조사 결과 중 검색 로봇이 측정할 수 있는 요소들로

결정한다. 이러한 요소에 따라 사용자가 원하는 요소의 값을 입력하면, 사용자가 프로파일이 결정된다. 검색 엔진은 웹에 존재하는 상품과 정보를 찾고, 그 결과를 데이터베이스에 저장한다.

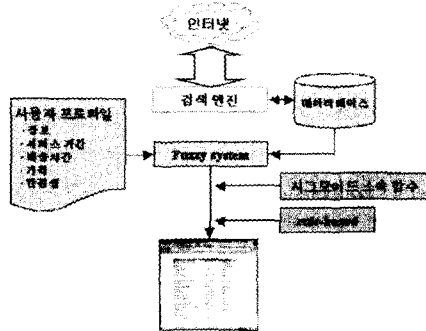


그림 1 시스템 구조

데이터베이스의 정보를 퍼지 기반의 시그모이드 소속 함수와 rule-base를 이용해서 사용자의 프로파일에 적합한 정보를 찾아 그 정보를 HTML 문서로 변환해서 사용자에게 검색 결과 리스트를 보여준다.

일반적인 검색 방법과는 달리 사용자 프로파일에 적합한 상품을 검색할 수 있는 방법을 제안하기 위해서 우리는 다음과 같은 세 가지 절차를 수행한다.

- i) 제품 선택 요소 결정, ii) 제품 선택 퍼지 시스템 생성, iii) rule-base 함수

3.1 제품 선택 요소 결정

전자 상거래에서 소비자가 상품을 선택하는 요소들을 결정하기 위해서 인터넷에서 상품이나 서비스를 판매하는 사이트를 통해 구매 경험이 있는 인터넷 이용자들을 대상으로 설문조사를 실시하였다.

설문항목은 설문자 인적사항을 포함하여 총 65 개 항목으로 구성되어 있다. 소비자 만족 측정을 위한 항목 (53항목), 사이트 사용 관련 항목 (5항목), 인터넷 사용에 관한 항목 (3 항목), 인적사항 항목 (4항목)으로 구성되어 있다. 설문조사 방법은 전자 메일을 이용했다.

그 결과 상품 정보, 소비자 서비스, 구매결과 및 배송, 사이트 화면 디자인, 구매절차, 안전성, 배송 시간 및 요금, 결제수단, 사이트 사용 편의성, 부가정보서비스가 산출되었다. 이 결과 중에서 웹 검색 로봇에 의해서 측정할 수 있는 요소로는 상품 정보(Product information:Pi), 서비스 기간(Term of service:Ts), 배송 시간(Delivery Time:Dt), 안전성(safety:Sa), 가격(Price:Pr) 등이 있다.

3.2 제품 선택 퍼지 시스템 생성

사용자의 프로파일에 적합한 상품을 나타내기 위한 요소들의 퍼지 집합 A가 임의의 $X=\{x\}$ 에 대하여 $[0,1]$ 값으로 표현되기 위해서는 $x=x_0$ 에 대해 집합 A의 소속 정도(membership degree)를 나타내는 소속 함수, $\mu_A: X \rightarrow [0,1]$ 으로 표현할 수 있다.

퍼지 집합의 특성은 5개의 입력 요소들에 따라 복잡한

값으로 구성되며, 이러한 요소들은 퍼지에 따라 취급되고 모두 소속 정도가 $[0, 1]$ 의 내부의 값을 갖는다.

정의1) P_i, Ts, Dt, Sa, Pr, Ff (Fuzzy function) $\div [0, 1] \rightarrow [0, 1]$

정의2) $Ff \div f(P_i, Ts, Dt, Sa, Pr, Ff)$ 은 $Ff \div ([0, 1] \dots [0, 1]) \rightarrow [0, 1]$

퍼지 시스템은 정의 2와 같이 표현하고, 상품 시스템을 생성하기 위한 절차는 다음과 같다.

1) 요소에 대한 등급

소속 함수는 각 요소들은 5가지 등급으로 이루어진 보편적인 집합 내에 정의된다. 즉, very low, low, middle, high, very high로 각 요소는 $[T_1, T_2]$ 가 $[0, 1]$ 로 사상된다.

본 논문에서 퍼지 소속 함수로 정의한 시그모이드 소속 함수는 다음의 세 가지 특징을 만족한다. 입력 값에 무관하게 항상 $[0, 1]$ 사이의 퍼지 값을 갖고 S자 형태의 단조 증가 형태를 이룬다. 또한, 임계값(critical value)을 갖는 퍼지 소속 함수이다[7].

2) 사용자 프로파일에 따른 시그모이드 함수

α -cut은 소속 함수의 $[0, 1]$ 사이의 값에서 임의의 α ($0 \leq \alpha \leq 1$)값이 되는 함수 값에 대한 퍼지 상태 변수의 구간을 나타낸다. 이 α -cut은 퍼지 집합의 원소들에 대해 집합에 속할 기준을 정의할 때 사용된다.

설문 조사에 따른 제품 선택 항목에 대한 레벨을 사상시키기 위해서 그림 2와 같이 각기 다른 임계값을 갖는 시그모이드 소속 함수를 정의하고, 또한 이러한 임계값은 요소들에 대한 실제 값에 따라 α -cut의 α 값이 결정된다.

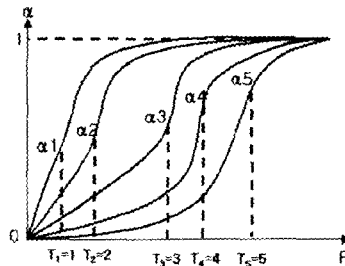


그림 2 등급에 대한 시그모이드 소속 함수 정의 임의의 X를 원소로 하는 퍼지 집합 A에 대해서 임의의 $\alpha \in [0, 1]$ 값을 가진 α -cut을 적용한 퍼지 집합 A_α 는 다음과 같이 정의한다.

정의3) $A_\alpha = \{x \mid A(x) \geq \alpha\}$

따라서, 퍼지 집합 A_α 는 퍼지 집합에 속할 소속 정도의 값이 α 값 이상으로 이루어진 집합이다. 본 논문에서는 사용하는 레벨에 대한 α -cut 값은 다음과 같다.

$$\alpha\text{-cut(very low)} = \text{Min_P_value}$$

$$\alpha\text{-cut(low)} = \left(\sum_{i=1}^n P_value/n \right) \times 0.2$$

$$\alpha\text{-cut(medium)} = \left(\sum_{i=1}^n P_value/n \right) \times 0.4$$

$$\alpha\text{-cut(high)} = \left(\sum_{i=1}^n P_value/n \right) \times 0.6$$

$$\alpha\text{-cut}(\text{very high}) = \left(\sum_{i=1}^n P\text{-value}/n \right) \times 0.8$$

3.3 사용자 프로파일에 적합한 검색 알고리즘
퍼지 개념을 이용한 사용자 프로파일에 따른 최적의 상품을 검색하는 알고리즘은 다음과 같다.

```
1) 5가지 요소의 등급 확인
입력 = factor_F[i] : 각 요소의 실제 값
출력 = 각 제품들의 요소에 대한 등급
function L_Check() {
  query("select factor_F[i] from Product");
  while (rs.next()) {
    if (factor_F[i] >= (alpha-cut[i](very high)))
      return "very high";
    else if (factor_F[i] >= (alpha-cut[i](high)))
      return "high";
    else if (factor_F[i] >= (alpha-cut[i](middle)))
      return "middle";
    else if (factor_F[i] >= (alpha-cut[low]))
      return "low";
    else
      return "very low";
  }
}
```

2) 사용자 프로파일에 따른 상품의 유사성 측정
입력 : 사용자 프로파일
출력 : 등급의 일치 개수
//Distance() : 사용자 프로파일과의 거리

```
function C_Number() {
  preference = 0;
  for (int j=0; j<=4; j++) {
    if (Distance(0))
      preference = preference + 1;
    else if (Distance(1))
      preference = preference + 0.6;
    else (Distance(1))
      preference = preference + 0.2;
  }
  Value = preference/5;
}
```

4. 시스템 설계

시스템은 사용자가 쉽게 사용할 수 있는 그래픽 사용자 인터페이스와 인터넷 환경에서 사용할 수 있도록 설계하였다. 그림 3은 사용자가 원하는 물건에 대해 사용자의 프로파일을 입력하는 화면이다.

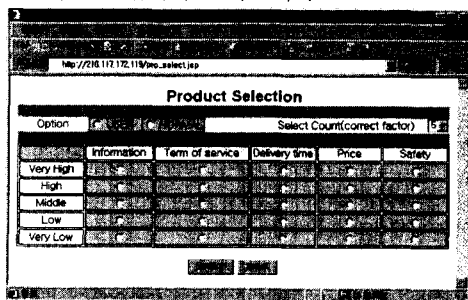


그림 3 상품 검색 사이트

모든 구현은 자바를 기반으로 둔 JSP(Java Server Page)와 서블릿(Servlet)을 사용하였고, 사용자의 편의를

위해서 제품 구입 요소들의 값을 입력하는 형태를 두 가지로 주었다. 하나는 가장 일반적인 키워드 검색을 이용한 것이고, 다른 하나는 퍼지 시스템을 이용한 것이다. 퍼지 시스템을 이용할 때는 검색에 대한 임계값을 사용자가 직접 입력할 수 있게 설계했다.

5. 결론 및 향후 연구

전자 상거래를 활성화하기 위한 여러 가지 연구들이 진행되었다. 전자 상거래에서 물건을 검색하기 위한 질의에 대한 처리 속도와 그 결과도 향상되었으나, 소비자는 요구 사항에 적합한 물건을 찾기 위해서 많은 시간과 피드백 과정을 거쳐야 한다.

본 논문에서는 소비자의 물건 구입에 영향을 미치는 요소들을 전자 메일을 통한 설문 조사로 추출하고, 이러한 요소들 중에서 검색 로봇에 의해서 측정될 수 있는 요소들을 추출하였다.

그리고 사용자가 원하는 제품을 검색 엔진을 이용해서 검색하고, 검색된 정보 가운데서 사용자 프로파일에 적합한 제품들을 퍼지이론을 적용하여 선별하는 전자상거래 검색 시스템을 설계하였다. 본 논문에서 제안한 전자상거래 검색 시스템은 사용자의 요구 사항에 적합한 상품을 우선 순위로 보여주기 때문에 사용자가 원하는 상품을 여러 번의 피드백을 거치지 않고 빠른 시간 내에 검색할 수 있다.

향후 연구로 본 논문에서 제안한 시스템을 구현하고, 실험을 통해 성능을 평가할 것이다.

참고 문헌

[1] Hye-Jue Eun, Yan Ha, Yong-Sung Kim, László T. Kóczy, User-query extension and Documents Categorization Based on fuzzy relations and fuzzy functions, Bioprocess System. 2000. 9.
 [2] Ji-Hyu Jeng, Extended Electronic Catalog For Dynamic and Flexible Electronic Commerce, 2000. 2
 [3] George J. Klir / Bo Yuan, Fuzzy Sets and Fuzzy Logic , 1998
 [4] K. Chakrabarty, L. T. Kóczy and T. D. Gedeon: Information retrieval in legal documents by fuzzy relational charts, submitted to JASIS.
 [5] M. R. Garey and D. S. Johnson: Computers and Intractability. A Guide to the Theory of NP-Completeness, W. H. Freeman and Co., San Francisco, 1979.
 [6] L. T. Kóczy and T. D. Gedeon: Information retrieval by fuzzy relations and hierarchical co-occurrence, Part I, IETR 97-01, Information Engineering, School of Computer Science, University of New South Wales, Sydney, 1997.
 [7] L. T. Kóczy and T. D. Gedeon: Information retrieval by fuzzy relations and hierarchical co-occurrence, Part II, IETR 97-03, Information Engineering, School of Computer Science, University of New South Wales, Sydney, 1997.