퍼지 및 지능시스템학회 논문지 2007, Vol. 17, No. 6, pp. 754-759

부정적 피드백과 강화학습을 이용한 사용자 프로파일 학습

Learning for User Profile Based on Negative Feedback and Reinforcement Learning

손기준 · 임수연 · 이상조
Ki-Jun Son, Soo-Yeon Lim and Sang-Jo Lee

경북대학교 컴퓨터공학과

요 약

정보추천은 동적인 사용자의 정보요구를 기준으로 문서들을 선별하여 제시한다. 사용자의 정보요구는 하나 이상의 단어들로 구성된 프로파일로 표현이 되며, 문서의 추천 과정 동안에 발생하는 사용자의 관련성 평가를 통해 구체적인 내용으로 변할 수 있다. 기존 연구의 경우 사용자는 자신이 직접 관련성 평가에 참여하여 평가 정보를 입력하고, 사용자가 평가한 긍정적 피드백 정보를 이용하여 사용자 프로파일을 학습한다. 본 연구는 사용자가 평가한 긍정적 피드백, 부정적 피드백 정보와 강화학습을 이용한 사용자 프로파일의 학습 방법을 제안한다. 제안된 방법과 기존의 긍정적 평가 정보만을 이용한 방법의 성능을 측정하기 위해 동일한 토픽에 대한 추천 성능을 비교 실험하였다. 실험한 결과 제안된 방식에 의한 추천 성능이 긍정적 평가 정보만을 이용한 피드백 방법보다 나은 성능을 보였다.

키워드 : 프로파일 학습. 관련성 피드백, 부정적 피드백, 추천 시스템.

Abstract

The information recommendation system offers selected documents according to information needs of dynamic users. User's needs are expressed as profiles consisting of one or more words and may be changed into some specifics through relevance feedback made by users during the recommendation process. In previous research, users have entered relevance information by taking part in explicit relevance feedbacks and learned user profiles using the positive relevance feedbacks. In this paper, we learn user profiles using not only positive relevance feedback but negative relevance feedback and reinforcement learning. To compare the proposed with previous method, we performed experiments to evaluate recommendation performance of the same topic. As a result, the former shows the improved performance than the latter does.

Key Words : Profile Learning, Relevance Feedback, Negative Feedback, Recommender System.

1. 서 론

인터넷의 발달로 인하여 전자 정보의 양이 증가함에 따라 사용자의 정보요구에 보다 적합한 정보를 찾는 요구가 비례적으로 늘어가고 있다. 사용자는 정보의 양이 증가할수록 모든 정보를 확인해 가며 자신이 원하는 정보를 찾아보기가 힘들어지게 된다. 따라서 컴퓨터가 사용자를 대신하여 양질의 정보를 여과해주는 여과 시스템은 매우 유용한 도구가 될 것이다.

여과 시스템은 이러한 방대한 정보 집합 속에서 사용자의 정보요구와 선호도에 대한 관련정보를 제시함으로써 사용자를 돕도록 의도 되어진 지능적인 시스템이라 정의할 수 있다. 개인화된 지능적 정보 에이전트는 사용자의 정보요구와 선호도를 직접적 또는 간접적으로 학습하여 사용자 프로파일을 구축하게 된다[1].

사용자의 선호도를 학습하기 위해 사용자의 평가를 이용한 피드백 방법을 이용한다. 이러한 피드백 방법(relevance feedback)은 사용자로부터 주어진 명시적(explicit)인 평가값을 이용하여 프로파일을 학습하며, 사용자 프로파일이 생성되면 이 프로파일을 이용하여 정보를 여과한다. 이를 이용한 시스템으로는 SIFT[2], LIRA[3], Ringo[4], Syskill & Webert[5]등이 있다. 이러한 시스템들은 사용자가 평가한 긍정적(positive) 관련성 정보를 이용하여 사용자 프로파일을 학습한다.

본 논문은 사용자의 선호도 학습을 위해 사용자가 평가한 긍적적 평가 정보와 부정적(negative) 평가 정보를 함께 이용하여 선호도로 부여하고, 사용자 선호도 학습에는 강화학습을 이용한다. 이렇게 학습된 프로파일을 이용하여 사용자의 선호도에 적합한 문서를 여과하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 관련성 정보를 이용한 사용자 프로파일 학습과 강화학습에 대하여 알아보고, 3장에서는 사용자 평가 정보와 강화학습을 이용한 사

접수일자 : 2006년 7월 19일
완료일자 : 2006년 11월 24일

용자 선호도 학습방법에 대하여 살펴본다. 4장에서는 부정적인 평가 정보와 이를 이용한 문서 여과 시스템에 대하여 살펴보고, 5장에서는 실험을 통하여 제안한 방법의 성능을 평가하고, 마지막으로 결론 및 향후 연구에 대하여 토론한다.

2. 관련연구

2.1 관련성 정보를 이용한 프로파일 학습

사용자 프로파일의 학습을 위해 사용자로부터 명시적으로 평가를 받는 방법으로는 SIFT[2], LIRA[3], Ringo[4], Syskill & Webert[5] 등이 있다. SIFT[2]는 단순한 단어 목록으로 이루어진 프로파일을 사용한다. 사용자는 자신의 관심분야 마다 각각의 프로파일을 만들어야 하고 사용자가 직접 수정하거나 적합성 피드백을 이용하여 시스템이 자동으로 수정한다.

LIRA[3]는 하루에 10개씩 새로운 URL을 여과하고 사용자에게 URL과 연결된 문서에 대한 평가를 5점에서 -5점까지 점수로 받고 이렇게 받은 점수에 비례하여 그 문서에 대한 가중치를 준다. Ringo[4]는 사용자로부터 음악 앨범과 가수에 대한 여과를 하는 시스템으로 사용자로부터 1에서 7점까지 점수를 받고 이를 피드백으로 사용하며, 사용자 사이의 유사도를 결정하는데 이용한다. Syskill & Webert[5]는 웹페이지에 대한 사용자의 평가정보를 사용하여 프로파일을 학습한다.

위에서 기술한 방법은 여과 과정 동안 발생하는 사용자의 관련성 평가 정보를 이용하여 추상적인 초기 정보 요구를 보다 구체적으로 표현하는 방법들이다. 이러한 방법들은 사용자의 피드백을 이용하여 프로파일을 학습하고, 이를 여과에 이용한다.

2.2 강화학습

학습이란 과거의 경험을 이용하여 현재의 문제를 해결하기 위한 지식이나 기술을 획득하고 개선하는 것을 의미한다. 이와 같은 학습은 어떻게 훈련하느냐에 따라 분류(classification)와 예측(prediction) 등과 같이 명시적으로 학습 목표가 주어지는 감독학습(supervisor learning), 군집화(clustering)와 같은 무감독학습(unsupervisor learning) 그리고 강화학습(reinforcement learning)으로 분류된다[1].

감독, 무감독학습 방법은 정형화된 사례나 지식으로부터 학습을 통해 주어진 환경을 학습하는 형태로 정형화가 가능한 정적(static)인 환경에 주로 응용되며, 특정 시스템의 동적인 상태를 표현해야 하는 응용에는 적합하지 못하다. 대부분의 실제 응용은 이러한 동적인 환경과 각 학습 시스템과의 상호작용이 고려되어야 하는 특성을 가지고 있는데, 강화학습은 동적인 환경에서 시행착오를 거치면서 환경으로부터 주어진 보상을 최대화 하기위한 학습방법으로 그 특성상 이러한 응용에 적합하다.

강화학습은 감독학습과 무감독학습의 중간적인 특성을 띠고 있다. 학습의 주체인 에이전트는 환경의 상태를 관측하고 과거의 경험을 바탕으로 행동을 선택하며, 이에 따른 보상을 환경으로부터 받게 된다. 즉, 강화학습은 동적인 환경에서 시행착오(trial-and-error)를 거치면서 환경으로부터 주어진 보상을 최대화 하기위한 학습 방법이다[6,7].

여과 시스템은 이러한 동적인 상태 즉, 여과 대상이 되는 문서가 없어지거나 관련 문서가 변경된 경우와 사용자의 정보 욕구의 변화에 따른 사용자 프로파일의 변화 등의 동적인 상황이 많이 존재한다. 따라서 여과 대상 문서를 수집하고 여과하는데 강화학습을 이용할 수 있다.

강화학습은 학습자가 주어진 환경과의 상호작용을 통해 자신의 학습 목표를 달성하는 실시간 학습으로 크게 다음의 네 가지 요소로 구성되어 있다.

2.2.1 환경 모델

환경(environment)은 주로 상태로 표현되며, 행위자(agent)는 적절한 정책(policy)에 따라 행동(action)을 취하게 된다. 이때 환경은 행위자에게 행동에 대한 보상(reward)을 주게 된다. 강화학습 에이전트가 t시각에 행동 a_t을 취하면 행동에 대한 보상 r_t가 환경으로부터 주어진다. 그리고 행동에 의하여 상태 s_t가 s_{t+1}로 변화된다. 이러한 행위자와 환경과의 상호작용은 그림 1과 같다.

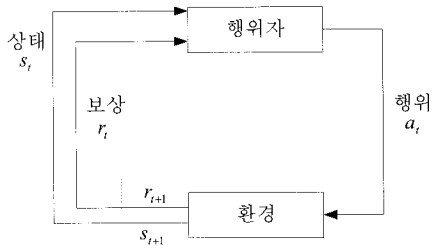


그림 1. 행위자와 환경간의 상호작용
Figure 1. The agent-environment interaction in reinforcement learning

2.2.2 정책

정책은 강화학습의 행위자가 특정 시간에 주어진 상태 집합에서 특정 상태를 선택하는 행동의 기준을 말하며, 확률적인 시행을 되풀이하여 이를 통해 얻어진 경험을 통해 얻어진다.

주어진 상태 집합에서 행동을 선택하는 방법은 세 가지 탐욕적 행동 선택 정책(greedy action selection), ϵ-탐욕 행동 선택 정책(ϵ-greedy action selection), 소프트맥스 행동 선택 정책(softmax action selection)으로 나눌 수 있다[6].

본 연구에서는 탐욕적 행동 선택 정책을 사용한다. 탐욕적 행동 선택 정책은 주어진 시간 t에 추정된 행동의 값들이 가장 높은 값을 갖는 행동을 선택하는 방법으로 행위자는 현재의 지식만을 이용하게 된다. 이 방법은 식 (1)과 같다.

$$Q_t(a^*) = \max_a Q_t(a) \tag{1}$$

위 식에서 a^*는 주어진 시간 t에 최적의 행동이다.

2.2.3 보상함수

보상함수(reward function)는 행위자의 행동에 대한 환경으로부터의 반응으로 보통 스칼라(scalar)값으로 주어진다. 강화학습에서 주어지는 보상은 무감독학습에서처럼 지시적인(indicative) 특성을 갖는 것이 아니라 평가적인(evaluative) 특성을 가지며 식 (2)과 같다.

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \tag{2}$$

위 식에서 $\gamma(0 \le \gamma < 1)$는 할인상수(discount rate)로 미

래에 대한보상을 현재 상태의 가치나 상태-행동의 가치에 반영되는 정도를 조절하며, t시간 이후 정책이 현재 정책에 반영되는 정도를 표시한다. 즉 γ의 값이 1에 가까울수록 t시각 이후에 받게 될 보상을 할인하지 않고 반영하게 된다.

2.2.4 가치함수

가치함수(value function)는 학습이 계속되는 동안 누적되는 보상과 현재의 상태에 대한 가치를 평가하는 함수이다. 일반적인 가치함수는 식 (3)과 같다.

$$V(s) \leftarrow V(s) + \alpha[V(s') - V(s)] \quad (3)$$

가치함수는 현재 상태(상태-행동의 쌍)의 가치를 평가하여 다음 상태의 가치를 추정하여 학습하는 규칙이다. 여기에서 α는 현재 상태의 변화량, 학습량(learning rate)을 조절하는 상수이다. 학습 목표에 가깝게 추정된 값 $V(s')$와 현재 상태의 값의 차이를 α만큼씩 현재의 상태를 변화시켜 가는 형식으로 학습 목표를 달성한다.

2.3 연관성 평가

사용자 프로파일과 여과된 문서간의 유사도를 측정하기 위해 여과 대상 문서를 tf(term frequency)벡터로 변환하고, 용어 빈도 순위가 높은 n개로 특징 벡터를 구성한다. 유사도의 정량화는 두 벡터간의 코사인 유사도 측정법 식 (4)를 사용한다[8].

$$sim(d_j, p_k) = \frac{\vec{d_j} \times \vec{p_k}}{|\vec{d_j}| \times |\vec{p_k}|} = \frac{\sum_{i=1}^{n} w_{i,j} \times w_{i,k}}{\sqrt{\sum_{i=1}^{n} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{n} w_{i,k}^2}} \quad (4)$$

식 (4)에서 p_k는 주제어 k에 대한 사용자 프로파일의 벡터이며, d_j는 문서 j의 tf벡터이다.

3. 사용자의 평가정보를 이용한 선호도 학습

지능적인 정보 에이전트의 사용자 프로파일은 사용자의 선호도를 해석하고 정보 여과 과정에서 이를 충분히 반영하는데 사용된다. 즉 사용자에게 유익한 관련 정보만을 선별적으로 제공하여야 한다. 이러한 특징을 지니기 위해서는 무엇보다도 사용자에 대한 프로파일의 정확한 구축이 필수 조건이며, 이를 위해 사용자의 평가 정보가 필요하다.

3.1 사용자의 평가 정보를 이용한 프로파일 학습

정보검색의 대표적인 모델인 벡터 공간 모델에서 사용자의 피드백을 평가할 때 사용자가 명시적으로 제공한 관련정보만을 이용하여 사용자의 프로파일을 학습하는 방법을 사용한다[8]. 이때 사용자가 관심을 가지지 않는 비관련 정보를 함께 이용함으로써 사용자의 관심 영역을 명확히 구분할 수 있다.

본 논문에서 학습 에이전트의 상태 s_t는 사용자 프로파일에 있는 용어들이며, 행동 a_t는 문서에 대한 사용자의 평가 정보에 의하여 용어를 재구성하는 것으로 사용자의 프로파일은 식 (5)에 따라 갱신된다.

$$w_{pk} \leftarrow w_{pk} + r_i \quad (5)$$

식 (5)에서 r_i는 여과된 문서 i에 대한 사용자의 관련성 평가 값이다. w_{pk}는 프로파일 p의 k번째 단어의 가중치이다. 이때 사용자가 명시적으로 제공한 정보에 대하여 긍정적인 평가와 부정적인 평가로 나누어 사용자의 프로파일 학습에 이용한다. 긍정적인 평가는 0~1 사이의 값으로 5단계로 나누어 사용하며, 부정적인 평가는 -1~-3의 3단계로 나누어 사용한다.

3.2 강화학습을 이용한 선호도 학습

선호도 학습은 사용자의 평가만이 수치적으로 주어지는 환경에서 평가를 최대화한다는 목표를 가짐으로써 강화학습 모델을 적용할 수 있다.

학습 에이전트의 목표는 사용자가 관심을 보인 문서에 가중치를 부여하여 용어를 선별한 후 최적의 프로파일을 유지하는 것이다.

사용자로부터 주어지는 문서 D_i에 대한 관련성 피드백은 스칼라 값을 가지며 식 (6)에 의해서 구해진다.

$$R(i) = \sum f_v(i) \quad (6)$$

위의 식에서 f_v는 사용자의 긍정적, 부정적 평가에 따라 주어지는 가중치이다. 여과된 문서에 대한 사용자의 선호도가 프로파일의 단어에 직접 반영되도록 하였다. 즉, 프로파일 p의 k번째 단어가 적합성 평가를 받은 문서 D_i 내에 있으면 가중치 w_{pk}는 문서의 적합성 평가를 α만큼 조절하여 반영한다. α는 즉각적인 보상 값이 아닌 지연된 값들을 결정하는 할인상수(discount factor)이다. 즉 현재 즉각적으로 받는 보상 값의 할인율은 0.9로서 이는 멀리 있는 보상 값보다 가까이 있는 보상 값을 더 선호한다는 것을 나타낸다.

사용자의 프로파일은 식 (2)에서 주어지는 명시적인 피드백 값 r_i을 이용하여 식 (7)과 같이 사용자 프로파일을 갱신한다.

$$w_{pk} = w_{pk} + \alpha r_i \quad (7)$$

사용자 선호도 학습 과정은 사용자로부터 관련성 피드백을 받고, 이 정보를 이용하여 사용자 프로파일을 학습한다. 즉, 사용자의 관련성 피드백을 이용하여 프로파일을 수정하며, 사용자의 선호도를 잘 반영하게 하는 것이 프로파일 학습의 목적이 된다. 사용자로부터 주어진 긍정적, 부정적 평가 정보를 이용한 프로파일의 학습 과정은 그림 2로 요약할 수 있다.

1. 사용자 프로파일 p을 초기화. set $t \leftarrow 0$
2. 사용자의 명시적 평가.
   - 2.1 $s \leftarrow$ 현재 상태.
   - 2.2 ε-greedy 정책에 의해 행동 선택.
   - 2.3 사용자로부터 긍정적, 부정적 평가 피드백을 받음.
   - 2.4 보상값을 식 (6)로 계산.
   - 2.5 용어 가중치 변경.
5. 사용자 프로파일 갱신. 식 (7).
6. $t \leftarrow t+1$, 단계 2로 이동.

그림 2. 사용자 프로파일 학습
Figure 2. Learning of user profile

사용자의 프로파일은 하나 이상의 주제(Topic)와 해당 주제어를 포함하는 단어들의 가중치로 구성된다. 사용자 프로

파일 w_p는 주제어 p에 대한 프로파일로 식 (8)과 같이 하나 이상의 단어들과 각 단어의 가중치 벡터로 표현한다. 사용자의 정보요구를 표현하는 프로파일 벡터는 다음과 같이 표현된다.

$$w_p = (w_{p1}, w_{p2}, ..., w_{pk}, ..., w_{pn}) \quad (8)$$

위 식에서 w_{pk}는 단어 k의 가중치이며 사용자의 프로파일을 표현하기 위해 n개의 단어를 이용한다.

학습과정에서 프로파일의 용어 재구성은 ε-greedy 정책을 사용하여 선택한다. 프로파일의 n개의 용어들은 사용자가 관심을 보인 문서의 용어벡터와 프로파일 벡터를 병합한 용어들에서 가중치가 높은 n개로 구성한다. 프로파일의 각 용어들의 가중치는 사용자의 관심도를 학습한 결과를 반영한다. 생성되는 용어벡터의 가중치는 tf 값을 사용한다.

4. 사용자의 평가 정보를 이용한 신문기사 여과 시스템

본 연구에서 제안한 사용자 프로파일 학습 방법을 이용한 여과 시스템은 그림 3과 같다. 여과 시스템은 두 가지 여과 에이전트와 인터페이스 에이전트로 구성되게 된다. 다음절에서는 각 에이전트에 기능에 대하여 기술한다.

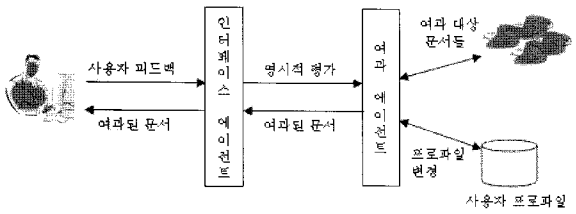


그림 3. 사용자 평가 정보를 이용한 문서 여과 시스템
Figure 3. Document recommender system using user rated feedback

4.1 여과 에이전트

여과 에이전트는 사용자를 대신하여 신문기사들 중에서 사용자에게 필요한 정보를 수집하고, 사용자의 선호도에 맞는 정보만을 선별하여 사용자에게 여과하는 역할을 한다. 이러한 역할을 수행하기 위해서 사용자로부터 명시적 평가를 받고 이를 이용하여 사용자의 프로파일을 생성한 후 사용자의 관심에 해당하는 신문기사를 여과한다.

여과 에이전트는 사용자가 선호한 문서로부터 선호도를 파악하여 사용자 프로파일을 구성하고, 사용자의 프로파일과 문서의 유사도를 결정하여 유사도가 높은 문서를 사용자에게 여과한다. 이렇게 여과된 정보로부터 사용자의 긍정적, 부정적 평가 정보와 강화학습을 이용하여 프로파일을 갱신하고 유지하게 된다.

4.2 인터페이스 에이전트

인터페이스 에이전트는 여과된 문서에 대한 사용자의 직접적 평가를 피드백 값으로 받아 이를 통합하는 역할을 한다. 이러한 피드백 값은 다시 여과 에이전트의 입력으로 들어가 사용자의 프로파일을 갱신시키는데 이용되게 된다. 이 사용자 프로파일을 이용하여 여과 에이전트는 높은 보상값을 받은 문서를 여과하게 된다. 이 문서는 다시 사용자에게 보여지게 되어 평가를 받게 되고, 사용자의 정보요구와 선호도를 따라 학습하게 된다.

여과 시스템은 사용자로부터 피드백을 받아 프로파일을 구성하고, 사용자의 프로파일을 이용하여 사용자가 선호할 만한 문서를 여과해 주게 된다. 사용자의 관심도에 나타난 단어들의 $Q(a^*)$값을 이용하여 상위에 랭크된 문서를 여과함으로써 더 효율적인 여과를 할 수 있다. 이는 특정 도메인이라는 한정된 상황에서 성능에 많은 향상을 가져올 수 있다. 그러나 실제적으로 사용자의 프로파일은 사용자의 관심도의 변화와 시간의 경과에 따라서 달라질 수 있다. 본 연구에서는 사용자의 관심도 변화는 없다는 가정 하에 실험을 수행한다.

5. 실험 및 분석

5.1 실험의 목적

실험의 목적은 사용자의 명시적 평가 정보에서 긍정적 평가 정보만을 이용한 경우와 부정적 평가 정보를 함께 이용한 경우에 대하여 결과를 비교 한다. 실험은 다음 세 가지 경우에 대하여 수행하였다. 첫째, 강화학습을 이용한 경우, 둘째, 긍정적 평가 정보를 이용한 경우, 셋째, 긍정적 평가 정보와 부정적 평가 정보도 함께 이용한 경우에 대하여 실험을 수행한다.

5.2 실험 방법

실험에 사용된 데이터는 중앙일보 신문기사를 대상으로 수집된 총 400개의 문서를 사용한다. 이 문서의 구성은 중앙일보 웹 사이트에서 제공되는 10개의 카테고리에서 각각 40개의 문서를 수집 사용하였고, 문서의 내용은 정치, 경제, 스포츠, 문화 등으로 구성되어 있다.

여과 시스템의 성능을 평가하기 위해 네 가지 토픽에 대하여 여과 성능을 평가한다. 실험에 사용한 네 가지 토픽은 올림픽, 유가(油價) 기사, 연예계 기사, 문화 기사로 구성되어진다. 실험 조건상 사용자 관심도 변화는 없는 것으로 가정 하였다.

문서들을 전처리 과정을 통하여 TF(term frequency)벡터로 표현하고, 이 문서들에 나오는 단어들과 사용자의 프로파일에 나오는 단어들을 이용하여 보상함수 $R(s)$를 구하고, 사용자 프로파일과 문서간의 유사도를 측정하기 위해 코사인 유사도 측정법을 사용한다.

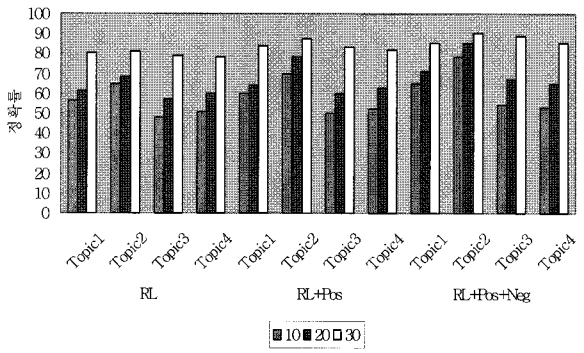
사용자의 정보요구와 선호도에 맞는 정보를 여과하기 위해 사용자 프로파일을 바탕으로 보상을 책정하는 것이 자연스럽게 이루어진다. 즉 현재의 문서가 얼마나 사용자 프로파일에 유사한지의 값을 바로 보상으로 줄 수가 있다.

5.3 실험 결과 및 분석

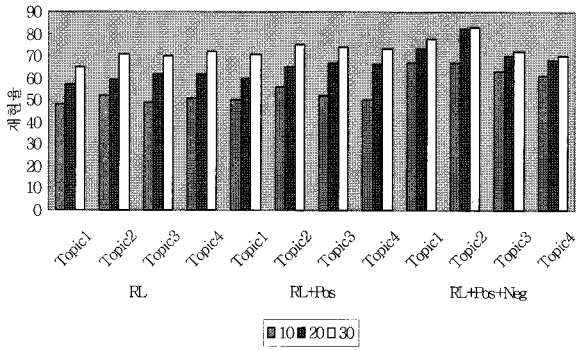
실험은 각 토픽별로 강화학습을 이용한 경우, 긍정적 평가 정보를 이용한 경우, 긍정적 평가 정보와 부정적 평가 정보를 함께 이용한 경우에 대하여 여과 시스템에 성능을 평가하였다. 여과 시스템의 성능 평가 척도로는 정확률과 재현율을 이용하여 평가 하였다.

평가 대상 문서로는 여과 시스템에서 여과되는 문서 중 상위에 랭크 된 10, 20, 30 개의 문서를 대상으로 한다. 제안한 방법을 여과 시스템에 적용한 실험결과는 그림 4와 같다. 그림 4에서 RL은 기준 시스템으로서 강화학습을 이용하여 여과한 결과이다. RL+Pos은 강화학습과 긍정적 평가 정보만

을 이용한 결과이며, RL+Pos+Neg은 강화학습과 긍정적, 부정적 평가정보도 함께 이용한 결과이다.



(a) RL, RL+Pos, RL+Pos+Neg의 정확률 비교
(a) Precision comparison of RL, RL+Pos and RL+Pos+Neg



(b) RL, RL+Pos, RL+Pos+Neg의 재현율 비교
(b) Recall comparison of RL, RL+Pos and RL+Pos+Neg

그림 4. RL, RL+Pos, RL+Pos+Neg의 실험결과
Figure 4. Experimentation result of RL, RL+Pos, RL+Pos+Neg

그림 4에서 볼 수 있듯이 사용자로부터 받은 긍정적 평가 정보만을 이용하여 프로파일을 학습 한 경우 보다 부정적 평가 정보도 함께 이용한 방법이 상위 10개의 문서를 대상으로 하였을 경우 정확률에서 각각 5%, 8%, 4%, 2%까지 향상 되었으며, 재현율은 각각 17%, 11%, 11.5%, 12%까지 향상 되었다.

또한, 여과되는 문서의 수가 많아지면 성능의 향상을 보이는 것을 알 수 있다. 여과 대상 문서를 상위 30개로 하였을 경우 정확률은 각각 82%, 90%, 89%, 85%로 나타났고, 재현율은 78%, 83%, 72%, 70%로 나타났다. 이 결과는 상위 20개의 문서를 대상으로 하였을 경우보다 정확률은 15%, 재현율은 9% 향상됨을 볼 수 있었다. 하지만 일반적인 주제를 다루는 Topic 3과 Topic 4는 큰 성능의 향상을 보이지 않았다. 또한, 그림 4(a)를 보면 Topic 2가 다른 토픽들 보다 더 나은 성능을 보이고 있다. 그 이유는 Topic 2에 등장한 어휘들이 다른 토픽에 나타난 어휘들 보다 분별력이 높기 때문으로 추정된다. 즉 그 주제에 나타난 어휘가 다른 주제에는 자주 사용되지 않기 때문이다. 예를 들면, 유가관련 문서에서 유가, 휘발유, OPEC, 두바이유와 같은 어휘들은 다른 주제에는 자주 사용 되지 않는 어휘이다. 따라서 이러한 어휘들이 유가관련 문서에 국한되어 나타남으로써 여과의 성능을 높여준다.

6. 결론

본 논문에서는 사용자의 긍정적, 부정적 평가 정보를 이용하여 사용자의 선호도를 학습하고, 이를 통하여 개인 사용자의 선호도에 적합한 문서를 여과하는 방법을 제시하였다. 기존의 수동적인 문서 여과 시스템은 많은 양의 문서들 중 사용자의 선호도에 맞는 문서를 선택해 주지만 동적인 환경에서 새로운 정보를 능동적으로 여과하지는 못한다.

본 연구에서 제안한 방법은 사용자의 긍정적, 부정적 평가 정보와 강화학습을 이용하여 사용자의 프로파일을 학습한 뒤 그 프로파일을 사용하여 사용자의 관심에 맞는 문서를 여과해주는 여과 시스템으로 동적인 환경에서도 능동적으로 사용자의 정보 요구에 응할 수 있는 장점을 가진다.

실험한 결과, 제시된 방법이 특정 주제에 대한 관심에 보다 적절한 문서들을 제시하는 것을 보았다. 즉 사용자가 제시한 긍정적, 부정적 평가 정보를 이용하여 사용자 프로파일을 학습하였을 경우 일반적인 긍정적 평가 정보만을 이용한 방법 보다 나은 성능을 보였다. 사용자들이 많은 양의 기사를 모두 읽고 자신의 관심 주제를 오류 없이 모두 명시하는 것이 쉽지 않은 일이므로, 어느 정도의 오류 및 누락과 관계없이 동작할 수 있는 이러한 특징은, 신문기사 여과 서비스와 같은 실용적인 용도의 여과를 위해 바람직한 것으로 보인다.

향후 과제로는 동적인 사용자의 관심변화 정보를 반영한 사용자 프로파일의 학습 방법에 대한 연구가 필요하다. 이와 더불어 문서에 대한 사용자의 선호도를 학습하는데 문서의 구조 정보를 이용한 방법과 문서의 연관성 평가 시 문서간의 구조적인 관계를 이용하는 방법에 대한 연구가 필요하다.

참 고 문 헌

[1] Seo, Y., Zang, B., "Personalized Web Document Filtering Using Reinforcement Learning," *Applied Artificial Intelligence*, Vol. 15(7), pp. 665-685, 2001.

[2] Tak W. Yan, Hector Garcia-Molina, "SIFT-A Tool for Wide-Area Information Dissemination," *Proceeding of the 1995 USENEX Techical Conference*, pp. 177-186, 1995.

[3] M. Balabanovic, Y. Shoham, "Learning Information Retrieval Agent: Experiments with Automated Web Browsing," In *Proceeding of the AAAI Spring Symposium on Information Gathering*, Stanford, CA, March 1995.

[4] Shardanand. U., and Maes "Social Information Filtering: Algorithmic for Automating Word of Mouth," In *Conference on Human Factors in Computing System(CHI'95)*, pp. 210-217, 1995.

[5] M. Pazzani, J. Muramatsu, D. Billsus "Syskill & Webert: Identifying interesting web sites," *National Conference on Artificial Intelligence*, vol. 1, pp. 54-61, 1996.

[6] R. S. Sutton, A. G. Barto, *Reinforcement Learning : An Introduction*, MIT Press, 1998.

[7] L. P. Kaelbling, M. L. L. Littman and A. W. Moore, "Reinforcement Learning: A Survey," *Journal of AI Research*, vol. 4, pp. 237-285, 1996.
[8] G. Salton, M. J. McGill, *Introduction to modern information retrieval*, McGraw Hill, 1983.
[9] 강승식, "HAM v.470c: 한국어 형태소 분석기와 한국어 분석 모듈," http://nlp.kookmin.ac.kr/ham/ham.html.
[10] T. M. Mitchell, *Machine Learning*, McGraw Hill , 1997.

저 자 소 개

손기준(KiJun Son)

제15권 1호 (2005년 2월호) 참조

임수연(SooYeon Lim)

제15권 1호 (2005년 2월호) 참조

이상조(SangJo Lee)

제15권 1호 (2005년 2월호) 참조