

수준별 프로그래밍 교육을 위한 단계별 클러스터링 기반 추천시스템

김 경 아*, 문 남 미**

The Recommendation System based on Staged Clustering for Leveled Programming Education

Kyung-Ah Kim*, Nammee Moon**

요 약

프로그래밍 교육은 학습자 개개인의 특성에 맞는 수준별·단계별 학습이 필요하다. 추천시스템은 개인화서비스를 위해 사용되는 방법의 하나로, 본 연구에서는 추천시스템을 사용하여 웹기반 프로그래밍 교육 환경에서 학습자 개개인에 적합한 학습을 추천할 수 있는 방법을 제공한다. 제안하는 수준별 프로그래밍 학습을 위한 추천시스템은 학습주제별 학습수준 기반 학습자 프로파일과 학습주제사이의 연관성 프로파일을 이용한 협업 필터링을 사용하여 특정 학습자의 학습수준과 학습범위에 적절한 프로그래밍 문제를 제공하도록 한다. 그 결과 프로그래밍 언어 교육 과정에서 발생하는 수준별·단계별 학습에 맞는 프로그래밍 문제 제공의 어려움을 해결하여, 학습자의 프로그래밍 능력 향상의 결과를 얻을 수 있었다. 더 나아가 기존 협업필터링 방법을 사용하는 경우와 비교해 볼 때 추천 성능 향상 및 분석 시간 감소를 통해 추천시스템의 한계점 중의 하나인 확장성을 해결할 수 있는 방법을 제시한다.

Abstract

Programming education needs learning which is adjusted individual learners' level of their learning abilities. Recommendation system is one way of implementing personalized service. In this research, we propose recommendation method which learning items are recommended for individual learners' learning in web-based programming education environment by. Our proposed system for leveled programming education provides appropriate programming problems for a certain learner in his learning level and learning scope employing collaborative filtering method using learners' profile of their level and correlation profile between learning topics. As a result, it resolves a problem that providing appropriate programming problems in learner's level, and we get a result that improving learner's programming ability. Furthermore, when we compared our proposed method and original collaborative filtering method, our proposed method provides the ways to solve the scalability which is one of the limitations in recommendation systems by improving recommendation performance and reducing analysis time.

· 제1저자 : 김경아 교신저자 : 문남미

· 투고일 : 2010. 07. 06, 심사일 : 2010. 07. 20, 게재확정일 : 2010. 08. 07.

* 명지전문대학 컴퓨터정보과 교수 ** 호서대학교 벤처전문대학원 IT응용기술학과 교수

- ▶ Keyword : 추천시스템(recommendation system), 협업필터링(collaborative filtering), 수준별 학습(leveled learning), 학습자 중심 학습(student-oriented learning), 클러스터링(clustering)

I. 서론

인터넷의 발달로 인해 많은 사람들은 웹을 친밀하고 익숙한 공간으로 인식하게 되었으며, 이러한 환경적 변화는 웹기반 교육 환경에서도 학습자들이 쉽게 적응하여 학습할 수 있는 기반을 조성해 주었다[1]. 웹기반 교육은 일반적으로 텍스트 정보와 함께 그림이나 사진 등 수업 내용과 관련된 시각 자료를 제공하고 동영상 및 음성파일도 삽입하여 학습 내용을 전달한다. 그러나 웹기반의 다양한 학습 자료는 교수자의 자료 게시와 일방적인 학습콘텐츠의 전달방식이 많았다[2].

웹기반 프로그래밍 교육을 위한 학습콘텐츠도 일반적으로 학습주제에 대한 강의노트와 이를 위한 다양한 예제와 설명 등으로 구성되어 있다. 물론 각 학습주제에 따른 프로그래밍 문제들도 제공된다. 프로그래밍 교육과정에서 프로그래밍 문제를 통해 문제 해결 능력을 향상하는 것은 무엇보다 중요한 학습방법이다[1]. 웹기반 프로그래밍 교육에서도 학습과정의 학습효과를 증진시키기 위해서는 학습자에게 적합한 프로그래밍 문제를 제공하는 것이 필요하다. 그러나 모든 학습자에게 일률적인 문제를 제공하는 것은 학습효과를 감소시킨다. 학습효과를 높이기 위해서는 학습자 개개인의 학습수준이나 학습내용 등에 따라 적합한 프로그래밍 문제를 제공하는 것이 필요하다. 따라서 프로그래밍 학습과정에서 학습자 개개인의 학습수준 및 학습단계에 맞는 프로그래밍 문제를 제공하는 수준별 학습은 프로그래밍 언어의 교육과정에서 매우 중요한 부분이다.

웹기반 프로그래밍 교육은 오프라인 프로그래밍 교육 환경보다 학습자의 학습자 중심 학습 방식이 필요한 교육 환경이다. 웹기반 프로그래밍 교육 환경에서 학습자는 학습과정에서 자신의 학습과정을 계획하고 학습자 자신에게 맞는 프로그래밍 학습과정과 프로그래밍 문제를 선택하고 구성할 수 있어야 한다. 그러나 현실적으로 다양하고 방대한 학습콘텐츠를 학습자 스스로가 계획하고 학습과정을 설계하는 것은 매우 힘든 작업이고[3], 학습자가 자신의 학습과정에서 학습효과를 증진시키는데 도움이 되는 학습자 수준에 적합한 프로그래밍 문제를 선택하는 데는 많은 어려움이 있다.

따라서 웹기반 프로그래밍 교육 시스템은 학습자가 학습과정에서 자신의 학습수준에 적합한 수준별 프로그래밍 학습을

할 수 있도록 도와주는 지원 시스템이 필요하다.

추천시스템은 특정 사용자에게 사용자가 선호할 것이라 예상되는 아이템들을 사용자에게 제공하는 시스템으로 대표적인 기법이 바로 협업필터링(collaborative filtering)이다. 협업필터링은 사용자기반 협업필터링과 아이템기반 협업필터링으로 구분된다. 사용자기반 협업필터링은 사용자의 선호 아이템을 예측하기 위해 과거의 평가 데이터를 분석하여 비슷한 선호도를 가지는 사용자들의 선호 아이템을 추천한다. 아이템기반 협업필터링은 사용자들이 과거에 좋아했던 아이템들과 비슷한 아이템은 좋아하고 싫어했던 아이템들은 싫어한다는 점을 기반으로 데이터를 분석하여 선호도를 예측한다[4].

프로그래밍 교육을 위한 이러한 콘텐츠와 같이 학습자의 학습단계, 선수지식상태, 학습수준별로 적합한 교육을 유도하기 위한 학습아이템을 추천하기 위해서는 기존 협업필터링 방법만으로는 맞춤형 학습을 위한 개인화 추천서비스에 한계가 발생한다. 교육적 측면의 추천시스템을 구축하기 위해서는 학습자가 선호하는 아이템의 추천보다 학습자의 학습단계에서 적합한 추천이 보다 의미 있는 추천이기 때문이다[5].

본 연구에서는 사용자기반 협업필터링을 보완한 혼합형 협업필터링 방식을 활용한 수준별 프로그래밍 교육을 위한 학습 지원 추천시스템을 제안하고, 기존의 협업필터링 방법과 비교 실험함으로써 제안된 추천시스템의 효율성을 증명한다. 제안된 시스템은 아이템간의 연관성을 고려하여 아이템들을 클러스터링하고, 사용자들 간의 아이템에 대한 이해도를 이용하여 사용자별로 최근접 이웃법에 의해 유사그룹을 구성하고 소속 그룹에 따라 사용자가 학습정보를 이용하여 아이템을 추천한다.

II. 관련 연구

1. 웹기반 프로그래밍 교육의 특성

1.1 수준별 학습

수준별 교육은 학습자의 흥미, 관심, 적성, 학습능력의 요구에 상응하는 교육의 내용, 방법, 기회를 제공하기 위해 도입된 것으로 단계형, 보충 및 심화형, 과목 선택형으로 분류하여 학습결손의 누적을 방지하기 위한 학습자 중심 교육의 실현이며 학교 교육의 질을 개선하기 위한 것이다[6].

프로그래밍 교육은 컴퓨터 교육의 어느 분야보다도 창의적인 학습자 중심의 개별 학습이 필요한 분야로 학습자 개개인의 수준, 학습 환경, 선행 학습 정도, 교과 수업의 참여도 등에 따라 학습의 결과가 현격히 차이가 난다. 또한 매 수업에 이루어지는 수업의 연속적인 특성으로 인해 단계별 학습의 완성수준이 단위 수업 전체의 결과에 많은 영향을 미친다[7,8,9].

따라서 효과적인 프로그래밍 언어 교육을 위해서는 학습자 개개인의 선행 학습 정도와 수준에 따라 개개의 학습동기 유발 방법 및 학습 이해도 향상을 고려한 수준별·단계별 교수 방법이 필요하다.

1.2 학습자 중심 학습

학습자 중심 학습모형이 적용되면서 학습자의 특성에 따라 학습전략이나 학습방법을 학습자가 스스로 선택할 수 있는 학습환경이 제시되고 학습자 개개인에 따라 학습내용, 학습방법 및 학습속도가 다르게 나타남을 알 수 있다[10]. 이는 교수자가 설정한 학습과정을 답습하는 일반적인 학습 전달 방식에서 학습자의 특성에 따라 학습자 스스로가 학습과정을 계획하고 실행하는 학습자 중심 학습의 필요성을 제기하는 계기가 되었다. 그러나 다양하고 방대한 학습콘텐츠를 학습자 스스로가 계획하고 학습과정을 설계하는 것은 매우 힘든 작업이다[3].

웹기반 프로그래밍 교육 환경에서 학습자는 학습을 계획하고, 계획에 따라 학습 문제를 탐구하고 해결해 가며 학습을 진행하게 된다. 이 과정에서 특히 프로그래밍 문제 중심의 학습 시 자신의 학습수준에 적합한 문제의 선택은 프로그래밍 학습과정에서 학습효과를 결정하는데 중요한 역할을 한다. 강의실에서 교수자와 함께하는 오프라인 수업에서는 수업과정에서 교수자가 학습자의 학습단계와 학습수준에 따라 단계별로 적절한 프로그래밍 문제를 제공하는 것이 가능하다. 그러나 웹기반의 온라인 교육 환경에서는 학습자가 스스로 자신의 학습단계와 학습수준에 적합한 프로그래밍 문제를 선택하여 학습하는 것이 쉬운 일은 아니다.

따라서 웹기반 프로그래밍 학습에 있어서 방대한 프로그래밍 문제로부터 자신의 학습에 적절한 프로그래밍 문제 선택을 학습자 중심 학습에서 할 수 있도록 지원하는 방법이 필요하다.

2. 협업필터링

추천시스템 연구의 주된 관심사는 고객과 아이템에 대한 이용 가능한 정보를 분석하여 고객들이 관심을 가질 아이템이 무엇인지를 파악하는 것이다. 추천시스템은 보통 아이템의 속성들을 기반으로 목표고객이 관심을 가진 아이템과 비슷한 속성의 아이템을 추천하거나 고객들이 아이템들을 경험하고 부여한 평가치들을 기반으로 목표고객이 평가하지 않은 아이템들

중에서 높게 평가할 것이라고 예상되는 아이템을 추천한다. 평가치는 고객들의 행동들을 주시하거나 고객들에게 경험한 아이템에 대한 평가치를 묻는 방법을 통해서 얻어진다[11].

협업필터링은 아이템에 대한 목표고객의 평가치와 다른 고객의 평가치를 이용하여 목표고객이 좋아할만한 아이템을 추천하는 기법이다. 확장성은 고객과 아이템의 개수가 늘어남에 따라 목표고객의 최근접 이웃을 찾기 위한 연산이 기하급수적으로 늘어나 추천목록을 생성하는데 까지 오랜 시간이 걸려 시스템의 효율성이 감소한다는 한계점이다[12]. 이는 전통적인 협업필터링 방법이 가지는 한계점 중의 하나로 여러 연구에서 확장성으로 인한 효율성의 문제를 개선하고자 하였다.

본 연구에서 제공한 수준별 프로그래밍 교육을 위한 단계별 클러스터링 기반의 추천시스템에서 활용한 협업필터링 방법은 프로그래밍 언어 교육에서 사용되는 학습주제들 간의 연관성 정도를 활용하는 방법이다. 이 방법은 기존의 협업필터링 방법을 적용한 경우와 비교할 때, 학습주제간 연관성 정도를 활용하여 학습주제와 관련된 아이템만을 대상으로 목표학습자의 최근접 이웃을 선정함으로써 협업필터링의 한계점 중의 하나인 확장성 문제를 개선할 수 있는 방법을 제공한다.

III. 본 론

1. 단계별 클러스터링 기반 추천시스템 개요

본 연구에서 제안하는 학습주제간 연관성을 고려한 협업필터링 추천시스템은 목표학습자의 유사학습자 그룹을 선정하는 과정에서 사용자기반 협업필터링을 보완한 혼합형 협업필터링 방식을 제안한다. 제안된 방법은 아이템간의 연관성을 고려하여 아이템들을 클러스터링하고, 사용자들 간의 아이템에 대한 이해도를 이용하여 사용자별로 최근접 이웃법에 의해 유사그룹을 구성하고 소속 그룹에 따라 사용자의 학습정보를 이용하여 아이템을 추천한다.

학습아이템을 클러스터링하는 과정에서 학습아이템 간의 연관성을 고려하여 클러스터링함으로써 사용자기반 협업필터링 과정에서 사용되는 학습아이템이 현재 학습주제와 연관성이 높은 학습아이템으로 구성되어 추천목록 생성 과정에서 학습자의 학습이해도가 고려된 추천이 가능하다. 즉 학습자의 학습수준에 적합한 난이도의 프로그래밍 문제 학습을 유도할 수 있다.

그림 1은 학습주제간 연관성을 고려한 클러스터링 기반 협업필터링 과정을 나타낸다.

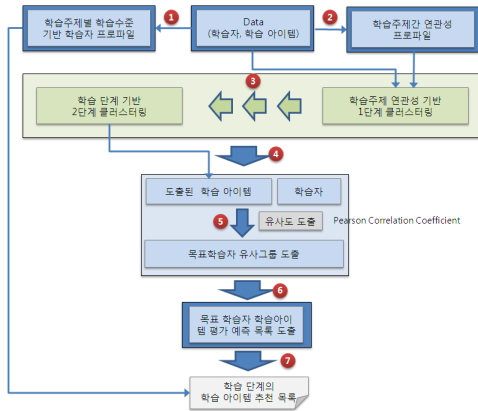


그림 1. 학습주제간 연관성을 고려한 클러스터링 기반 협업필터링 과정

Fig. 1. collaborative filtering process based on clustering using correlation between learning topics

① 학습자가 평가한 아이템을 바탕으로 아이템별 학습주제를 도출한 후 각 학습주제별 평균에 따른 학습정도로 학습주제별 학습수준 우선순위를 구해 학습수준 기반 학습자 프로파일을 생성한다. 학습자 u_i 의 각 학습주제 T_j 에 대한 학습수준 $(A_{i,j})$ 은 학습주제 T_j 를 포함하는 학습아이템의 학습결과 평점의 평균으로 구한다[5].

$$A_{i,j} = \frac{\sum_{k=1}^m S_k}{m} \dots\dots\dots (1)$$

식(1)에서 $S = \{ S_k : a \text{ score of } P_k \text{ which is chosen learning item in } L, 1 \leq k \leq m \}$
 $L = \{ L_j : W_{T_j} \} = 0.5 \text{ items}, 1 \leq j \leq n, n : \text{number of items}, W_{T_j} : \text{weight of } T_j, T_j : i^{th} \text{ learning topic}, L_j \in I$
 $I : a \text{ set of learning items}$ 를 의미한다.

② 학습주제들에 대해 전문가들로부터 연관성 정도값을 입력받아 평균값을 구해 학습주제간 연관도 프로파일을 생성한다. i 번째 학습주제와 j 번째 학습주제 간의 연관성 $R_{i,j}$ 는 각 전문가가 i 번째 학습주제와 j 번째 학습주제에 부여한 연관성 정도값의 평균값으로 구한다[5].

$$R_{i,j} = \frac{\sum_{k=1}^m R_{i,j,k}}{m} \dots\dots\dots (2)$$

식(2)에서 $R_{i,j,k} : \text{correlation value by } k^{th} \text{ expert for similarity between } i^{th} \text{ topic and } j^{th} \text{ topic}$
 $m : \text{number of experts}$ 를 의미한다.

③ 초기 데이터에 대한 아이템 속성(학습주제) 데이터를 수

집하고 분석하여 현재 목표학습자(u_a)의 학습주제와 학습주제간 연관도를 바탕으로 학습에 필요한 주제를 도출한 후, 이 연관 학습주제를 포함하고 있는 학습아이템의 집합을 도출한다. 이를 통해 하나의 프로그래밍 문제가 여러 학습주제들과 연관되어 있는 문제를 고려한 대상 아이템의 선정이 가능하다.

④ ③에서 도출된 학습아이템의 집합으로부터 현재 학습단계 보다 이후의 학습주제를 포함하고 있는 학습아이템을 제거한 학습아이템의 집합을 도출한다.

⑤ 목표학습자와 학습자들간의 유사도를 ④에서 생성된 아이템을 대상으로 식(3)의 피어슨상관계수(pearson correlation coefficient)[13]를 이용하여 계산한다.

$$P(u_a, u_i) = \frac{\sum_{j=1}^m (r_{a,j} - \bar{r}_a)(r_{i,j} - \bar{r}_i)}{\sqrt{\sum_{j=1}^m (r_{a,j} - \bar{r}_a)^2} \sqrt{\sum_{j=1}^m (r_{i,j} - \bar{r}_i)^2}} \dots\dots\dots (3)$$

식(3)에서 $r_{a,j} : \text{evaluated value of } u_a \text{ for learning item } j$
 $\bar{r}_a : \text{average of evaluated values of } u_a$
 $r_{i,j} : \text{average of evaluated values of } u_i$ 를 의미한다.

⑥ 다음, 최근접 유사그룹내의 유사도와 평가치를 이용하여 목표학습자가 학습하지 않은 아이템의 평가치를 예측한 후, 각 학습자의 학습결과치의 평균값과 유사도를 가중치로 적용하여 목표학습자(u_a)의 아이템에 대한 평가 예측치($P_{a,i}$)를 구한다[14]. $P_{a,i}$ 를 구하는 식은 식(4)과 같다. 평가 예측치의 값이 높은 학습아이템은 목표학습자의 학습결과가 높게 예측되는 학습아이템으로 학습 난이도 관점에서 보다 쉽게 해결할 수 있는, 학습 난이도가 낮은 학습아이템으로 예측된다.

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_a) \times P(u_a, u_i)}{\sum_{u=1}^n P(u_a, u_i)} \dots\dots\dots (4)$$

⑦ Top-N 기법으로 상위 N개의 추천 목록을 생성하여 제공한다. 학습자는 최소 1개 이상의 학습아이템을 학습한다. 학습한 평균학습결과가 기준 학습결과 점수를 통과하지 못하면 학습주제간 프로파일 정보를 활용하여 연관학습주제를 구한다. 다음 연관학습주제에 대한 학습자 프로파일을 활용한 확인을 통해 ⑥에서 도출된 연관학습주제와의 학습수준을 고려하여 연관학습주제의 비중이 최우선인 학습아이템을 구한다.

2. 시나리오 기반 서비스 실험

본 연구에서 제안한 수준별 프로그래밍 교육을 위한 추천

시스템에 대한 실험은 목표학습자에 대한 시나리오를 바탕으로 진행한다.

- 목표학습자는 현재 학습주제 For문을 학습하고 있는 학습자이다.
- 학습수준 전체 학습자의 동일 단계까지 학습한 학습자 중 상위 45%에 해당하는 성적을 보유하고 있다.

실험을 위하여, 먼저 전체 학습아이템을 대상으로 기존의 협업필터링을 적용하고, 다음으로는 본 연구에서 제안한 학습주제간의 연관성 정도를 고려한 학습아이템을 대상으로 협업필터링을 적용하여 상호 비교하고 분석한다.

본 연구에서 활용한 데이터의 집합은 본 연구자의 프로그래밍 수업을 수강한 431명의 학습자 중 온라인 학습에 참여한 168명의 학습자가 25개의 학습주제, 895개의 프로그래밍 문제에 대해 8,521건의 점수를 매긴 자료를 바탕으로 한다. 하나의 프로그래밍 문제는 최소 1개에서 최대 5개까지의 학습주제와 관련되어 있다.

첫 번째, 전체 학습아이템을 대상으로 한 모델은 목표학습자와 모든 학습아이템들에 대한 학습이해도가 유사한 그룹을 도출하고, 이들이 평가한 각 아이tem 평가치를 기반으로 유사도를 산출한다. 다음, 목표고객이 평가하지 않은 아이tem에 대한 평가치를 예측하고 Top-N기법으로 상위 5개 추천 목록을 생성하여 서비스할 수 있도록 한다. 표 1은 기존 협업필터링 방법을 사용한 모델에서 목표학습자에 대한 유사학습자 그룹의 유사도와 평가치를 이용해 목표학습자의 아이tem에 대한 평가 예측치를 도출하여 구한 5개의 추천결과를 나타낸다.

표 1. 기존 협업필터링 방법에 의한 추천결과
Table1. recommendation result using basic collaborative filtering

Student	Pearson	Score Mean	Each Programming Score				
			5	61	98	310	413
29	0.7354	80.0000	0.0000	0.0000	0.0000	0.0000	7.3535
19	0.7197	80.0000	0.0000	-7.1973	7.1973	7.1973	0.0000
49	0.7182	68.3333	-5.9846	1.1969	1.1969	1.1969	8.3785
93	0.7104	65.0000	-10.6563	-3.5521	3.5521	3.5521	10.6563
72	0.7083	86.6667	-4.7219	-4.7219	2.3609	2.3609	9.4438
64	0.7053	75.0000	-3.5263	-3.5263	3.5263	-3.5263	10.5789
83	0.7006	58.3333	-5.8386	1.1677	8.1741	1.1677	8.1741
111	0.6990	56.6667	-4.6597	2.3299	2.3299	2.3299	9.3194
122	0.6984	66.6667	-4.6557	2.3278	2.3278	2.3278	9.3113
120	0.6980	50.0000	-6.9805	0.0000	6.9805	6.9805	6.9805
Rating Sum			-47.0236	-11.9752	37.6458	23.5868	80.1963
Rating Sum/Pearson Sum			-5.1948	-1.3229	4.1588	2.6057	8.8594
Rediction Rate			64.8052	68.6771	74.1588	72.6057	78.8594
Rank			5	4	3	2	1

두 번째, 본 연구에서 제안한 모델은 목표학습자의 학습주제와 연관성 정도를 고려한 학습주제 연관성 프로파일을 기반으로 대상 학습아이템을 도출한다. 이를 바탕으로 학습자와 학습이해도가 유사한 유사그룹을 도출하고, 목표학습자(u_a)와 유사그룹내 학습자(u_i)의 유사도 $P(u_a, u_i)$ 를 피어슨상관계수를 이용하여 계산한 후, 유사도를 기준으로 유사그룹 학습자들을 정리한다. 다음, 이들의 평가치를 이용하여 목표학습자가 학습하지 않은 학습아이템의 학습 결과치를 예측하고 Top-N기법으로 상위 5개 추천목록을 생성하여 추천할 수 있도록 한다.

제안된 추천시스템 모델은 학습주제 연관성 정도값을 활용하여 대상 아이tem을 도출함으로써 기존 협업필터링 방법과 비교하여 현재 학습주제와의 연관성이 높은 학습아이템들만을 대상으로 유사학습자 그룹을 생성한다. 따라서 학습자의 현재 학습단계에서의 학습수준을 보다 정확히 예측함으로써 추천된 프로그래밍 문제의 난이도가 목표학습자의 현재 학습단계의 학습수준에 보다 적합하게 추천된다.

표 2는 본 연구에서 제안한 학습주제 연관성 기반 협업필터링 방법을 사용한 모델에서 목표학습자에 대한 유사학습자 그룹의 유사도와 평가치를 이용해 목표학습자의 아이tem에 대한 평가 예측치를 도출하여 구한 5개의 추천결과를 나타낸다.

표 2. 학습주제 연관성 기반 협업필터링을 사용한 추천결과
Table 2. recommendation result using collaborative filtering based on correlation between learning topics

Student	Pearson	Score Mean	Each Programming Score				
			5	61	89	102	137
21	0.9102	90.0000	-9.1020	0.0000	0.0000	9.1020	9.1020
35	0.9031	90.0000	0.0000	-9.0310	0.0000	9.0310	9.0310
49	0.9203	80.0000	0.0000	0.0000	0.0000	0.0000	9.2030
65	0.9341	81.6667	-1.5568	-1.5568	7.7842	-1.5568	7.7842
93	0.9341	80.0000	0.0000	-9.3410	9.3410	9.3410	0.0000
104	0.8680	90.0000	-17.3600	8.6800	0.0000	8.6800	8.6800
116	0.8805	68.3333	-7.3375	1.4675	1.4675	1.4675	10.2725
121	0.9062	78.3333	-7.5517	1.5103	1.5103	10.5723	1.5103
133	0.9210	68.3333	-7.6750	1.5350	1.5350	1.5350	10.7450
152	0.8746	60.0000	-8.7460	0.0000	8.7460	8.7460	8.7460
Rating Sum			-59.3290	-6.7360	30.3840	56.9180	75.0740
Rating Sum/Pearson Sum			-6.5542	-0.7441	3.3566	6.2878	8.2935
Rediction Rate			63.4458	69.2559	73.3566	76.2878	78.2935
Rank			5	4	3	2	1

3. 성능 측정

추천시스템의 예측 성능을 측정하는 지표로는 MAE(Mean Absolute Error), MAPE(Mean Absolute Percent Error),

RMSE(Root Mean Squared Error)가 보편적으로 이용된다[15]. 이 중 본 연구에서는 가장 많이 사용되는 MAE를 이용하여 고객이 실제 부여한 평가치(p_i)와 추천 알고리즘에 의해 예측된 평가치(r_i)의 차이로 성능을 측정한다.

$$MAE = \frac{1}{n} \sum_{i=1}^n \frac{|p_i - r_i|}{100} \dots\dots\dots (5)$$

MAE 지표는 예측된 평가치들이 실제 고객의 평가치들과 평균적으로 얼마나 흡사한지를 나타내는 지표로서, MAE 값이 작을수록 추천시스템의 예측 정확도가 높음을 의미한다.

식(5)은 본 연구에서 사용한 MAE 계산식으로 [15,16]에서 사용한 MAE 계산식에 기반하여 학습자 이해도 값을 보정하여 정의한다. 식(5)을 적용하여 III장 2절에서 기술한 두 개의 모델 별 실험을 통해 얻은 추천목록의 MAE를 구하면 표 3과 같다.

표 3. 모델별 성능 비교
Table 3. performance comparison depending on models

모델	전체 학습아이템을 대상으로 한 모델	학습주제 연관성을 고려한 학습아이템을 대상으로 한 모델
MAE	0.03631	0.03333

표 3에서 볼 수 있듯이 전체 학습아이템을 대상으로 한 기존 협업필터링 방법을 사용한 모델의 MAE 값은 0.03631이고 학습주제 연관성을 고려한 학습아이템을 대상으로 한 협업필터링 방법을 사용한 모델의 MAE 값은 0.03333이다. 이를 통해 추천을 위한 협업필터링 모델에 있어 학습주제 연관성을 고려한 학습아이템을 이용한 실험이 전체 학습아이템을 대상으로 한 모델 보다 평균 오차값이 감소한 것을 관찰할 수 있다.

평균 오차에서 볼 수 있듯이 학습주제 연관성을 고려한 학습아이템을 대상으로 한 모델의 예측률이 높을 뿐만 아니라 유사그룹 도출을 위한 대상 아이템 수를 감소시켜 협업필터링의 효율도 높일 수 있다.

그림 2와 같이 학습주제 연관성을 고려한 학습아이템을 선정하는 단계가 진행되면서 유사그룹을 구할 때 사용할 학습아이템의 수가 감소한다. 첫 번째 모델은 전체 학습아이템을 대상으로 유사그룹을 선정하였으므로 895개의 학습아이템을 대상으로 피어슨상관계수를 구하였다. 그러나 학습주제 연관성을 고려한 두 번째 모델은 전체 학습아이템 중 현재 목표학습자가 학습하고자 하는 단계에서 학습주제 연관성 프로파일의 정보를 기반으로 관련성이 높은 192개의 학습아이템을 대상으로 피어슨상관계수를 구하여 유사그룹을 선정하였다.

이 두 개의 실험 모델을 비교하면 두 번째 모델의 대상 아이템의 수가 첫 번째 모델의 대상 아이템의 수의 21.45%에 해당함을 알 수 있다. 이는 피어슨상관계수를 구하여 유사그룹을 선정하는 과정에서 피어슨상관계수를 계산하는 대상 아이템 수를 감소시켜 협업필터링 과정에 효율성을 제공한다.

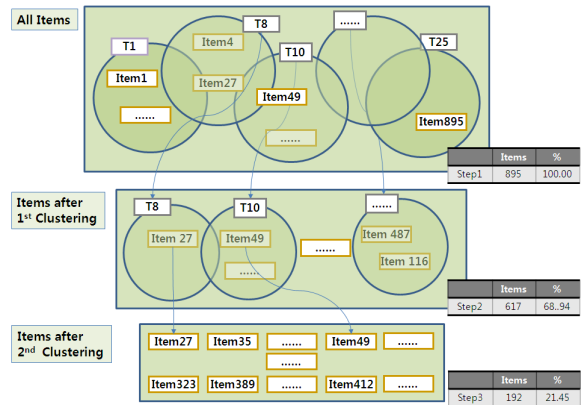


그림 2. 단계별 클러스터링을 통한 학습아이템 수 감소 흐름
Fig. 2. reduction flow in the number of learning items using staged clustering

표 4. 단계별 클러스터링을 통한 아이템수 감소율
Table 4. reduction ratio in number of items using staged clustering

	Step1	Step2	Step3
Items	895	617	192
비율(%)	100	68.94	21.45

결과적으로 본 연구에서 제안하는 수준별 프로그래밍 교육을 위한 단계별 클러스터링 기반 추천시스템에서는 목표학습자의 학습단계에서 학습자의 학습수준에 적합한 학습아이템의 추천 시 목표학습자의 학습단계와 관련성이 높은 학습아이템에 대한 학습자들 간의 유사도가 보다 정확하게 타겟팅된 추천목록을 도출할 수 있음을 의미한다. 이는 협업필터링의 성능 향상에 도움이 될 뿐만 아니라 목표학습자의 학습효과 측면에서도 의미 있는 추천결과를 도출한다.

IV. 결론

본 연구에서는 사용자기반 협업필터링을 보완하여 아이템 간의 연관성을 고려하여 아이템들을 클러스터링하고, 사용자

들 간의 아이টে에 대한 이해도를 이용하여 사용자별로 최근접 이웃법에 의해 유사그룹을 구성하고 소속 그룹에 따라 사용자가 학습정보를 이용하여 아이টে를 추천하는 방법인 혼합형 협업 필터링 방식을 제안하였다.

프로그래밍 언어 교육에서 사용하는 학습아이টে인 프로그래밍 문제 집합에서는 하나의 프로그래밍 문제가 하나의 학습 주제와만 관련되어 있는 것이 아니고 학습이라는 특성상 관련 학습주제에 대한 이해수준의 정도가 유사그룹을 선정하는데 의미 있는 속성이다. 따라서 본 연구에서 제안하는 학습주제 간 연관성으로 고려한 학습아이টে 클러스터링에 기반한 혼합형 필터링은 학습아이টে의 추천 과정에서 학습자의 이해수준과 프로그래밍 언어 교육의 학습적 특성을 반영한 학습자 중심의 학습을 유도할 수 있도록 지원한다.

학습주제별 연관성 정보를 이용한 단계별 클러스터링은 문제은행 방식의 많은 문제로부터 분석할 대상 아이টে 수를 감소시켜 분석 시간의 감소를 유도한다. 또한 학습자의 학습수준을 고려한 유사도 계산을 바탕으로 최근접 이웃학습자를 선정하고 유사그룹간의 정보를 통해 추천 성능을 높였다.

더 나아가 교육 서비스 부분에서 필요한 추천된 결과에 대한 학습자의 학습결과에 따른 피드백을 받아 학습자의 학습주제별 학습이해도를 기반으로 다음 추천에 적용할 수 있도록 연구함으로써 교육 서비스 부분에서 필요한 학습지원 추천시스템의 추천 방법을 제안하는데 또 다른 연구의 의의가 있다.

참고문헌

- [1] 김영지, 염용철, 김현철, 이원규, “웹기반 프로그래밍 교육 시스템의 설계 및 구현”, 한국컴퓨터종합학술대회 논문집, 제 32권, 제 1호, 67-69쪽, 2005년.
- [2] 정화영, 이연호, 홍봉화, “학습 선호도에 의한 학습 콘텐츠 제안 시스템”, 한국콘텐츠학회논문지, 제 10권, 제 1호, 477-485쪽, 2010년.
- [3] 정화영, 홍봉화, “자기주도적 학습을 위한 학습자 수준별 콘텐츠 구성”, 한국콘텐츠학회논문지, 제 9권, 제 7호, 402-410쪽, 2009년 7월.
- [4] 김청송, 이진승, 강재우, “향상된 아이টে 기반의 협동적 필터링”, 한국컴퓨터종합학술대회 논문집, 제 36권, 제 1호 (C), 469-473쪽, 2009년.
- [5] 김경아, 문남미, “프로그래밍 언어 학습지원 추천시스템”, 전자공학회논문지, 제 47권, CI편, 제 4호, 11-17쪽, 2010년. 7월.
- [6] 김명신, 김희숙, 정성태, “수준별 교육을 위한 웹기반 교수-학습 시스템의 설계 및 구현”, 한국정보과학회, 추계학술대회 논문집, 제 27권, 제 2호, 565-567쪽, 2000년.
- [7] 권재은, “프로그래밍 언어 수업을 위한 교수-학습 모형의 개발과 적용”, 석사학위논문, 안동대학교, 2004년.
- [8] 김경아, 김계현, 구윤모, “프로그래밍 언어 교육에 있어서 적용가능한 효과적인 교수-학습 방법에 대한 연구”, 명지전문대학 정보기술연구소논문집, 제 4권, 63-71쪽, 2004년.
- [9] 손경아, “문제중심학습(Problem-Based Learning)을 적용한 비주얼베이직(VisualBasic) 프로그래밍 학습의 사례연구”, 컴퓨터교육학회논문지, 제 5권, 제 1호, 17-25쪽, 2002년.
- [10] 박순일, 고병오, “전문자 학습활동 모형의 효율적 운영을 위한 웹 기반 교수-학습 시스템 개발”, 정보교육학회논문지, 제 8권, 제 3호, 293-306쪽, 2004년.
- [11] 이재석, 박석두, “장르별 협업필터링을 이용한 영화 추천시스템의 성능향상”, 한국지능정보시스템논문지, 제 13권, 제 4호, 65-78쪽, 2007년. 12월.
- [12] 이재석, 박석두, “단계적 협업필터링을 이용한 추천시스템의 성능 향상”, 한국지능정보시스템학회, 춘계학술대회논문집, 218-225쪽, 2007년. 5월.
- [13] 권병일, 문남미, “이러닝 마켓플레이스에서의 자기주도 학습지원을 위한 추천시스템”, 한국컴퓨터정보학회논문지, 제 15권, 제 2호, 135-146쪽, 2010년. 2월.
- [14] Badrul Sarwar, Geoege Karypis, Joseph Konstan and John Riedl, “Item-based Collaborative Filtering Recommendation Algorithm”, WWW10, pp.285-295
- [15] 김경록, 이주호, 변재희, 문남미, “모바일 환경에서 장르 속성을 이용한 추천 연구”, 한국멀티미디어학회, 춘계학술발표대회 논문집, 제 13권, 제 1호 2010년.
- [16] 이희춘, 김선옥, 이석준, “추천시스템에서 MAE와 표준편차의 상관성에 관한 연구”, 한국인터넷정보학회, 춘계학술발표대회논문집, 제 8권, 제 1호, 391-395쪽, 2007년.
- [17] 이민경, 강수용, “웹 기반 수준별 학습을 고려한 문제은행 시스템의 설계 및 구현”, 한국정보과학회, 가을 학술발표논문집, 제 33권, 제 2호(A), 103-107쪽, 2006년.
- [18] 홍태호, 이희정, 서보밀, “클러스터링 기반 사례기반추론을 이용한 웹 개인화 추천시스템”, 한국지능정보시스템학회논문지, 제 11권, 제 1호, 107-121쪽, 2005년 6월.
- [19] 김혜은, 유석중, “수준별 학습과 학습 관심도를 고려한 학습평가시스템”, 한국컴퓨터정보학회논문지, 제 13권, 제 6호, 69-76쪽, 2008년.

저 자 소 개



김 경 아

1990: 이화여자대학교 이학사.
1992: 이화여자대학교 이학석사.
2001: 이화여자대학교 공학박사.
2002 - 현재 : 명지전문대학 컴퓨터
정보과 부교수
관심분야 : 프로그래밍 언어, 이터닝,
추천시스템, HCI



문 남 미

1985: 이화여자대학교 이학사
1987: 이화여자대학교 이학석사
1998: 이화여자대학교 공학박사
2008 - 현재 : 호서대학교 벤처전문
대학원 IT응용기술
학과 교수
관심분야 : 디지털데이터방송비즈니스
모델, T-Commerce, 메
타데이터, HCI