

순차 연관 규칙을 이용한 개인화된 전시 부스 추천 방법

문현실

경희대학교 경영대학 & 경영연구원
(pahunter@khu.ac.kr)

정민규

경희대학교 경영대학 & 경영연구원
(minkyuli@khu.ac.kr)

김재경

경희대학교 경영대학 & 경영연구원
(jaek@khu.ac.kr)

김혜경

경희대학교 경영대학 & 경영연구원
(kimhk@khu.ac.kr)

.....

전시회는 전시업체가 새로운 상품이나 서비스를 관람객에게 알리기 위해 개최되는 것으로 효과적인 마케팅 수단으로 중요한 역할을 수행한다. 전시회를 방문하는 다양한 관람객의 니즈를 충족시키기 위하여 다양한 유비쿼터스 기술이 전시회에 응용되고 있지만 관람객이 사전에 요청한 정보만을 제공함으로써 개별 관람객의 선호가 반영되지 않아 관람객의 니즈를 충족시키기에는 한계가 있다. 이러한 한계를 해결하기 위한 방법으로 개인의 선호에 부합하는 부스를 추천하는 추천 시스템의 이용이 가능하다. 추천시스템은 전시 환경에서 관람객의 선호를 추론하여 선호에 부합하는 방문 부스를 추천하여 관람객의 니즈를 충족시킬 수 있다. 그러나 추천 시스템 중 가장 성공적으로 평가 받는 기존의 협업 필터링은 관람객의 부스 방문 순서에 나타나는 선호를 반영하지 않아 동적으로 변화하는 선호를 가지는 관람객으로 구성된 전시 환경의 추천 시스템으로는 적합하지 않다. 따라서 본 연구에서는 관람객의 방문 순서를 고려하는 기법 중 순차 연관 규칙을 이용하여 관람객의 선호에 부합하는 부스를 추천하는 방법론을 제안하였다. 본 연구에서 제안한 방법론의 성과 측정을 위해 실제 전시회에서 획득한 데이터를 사용하여 기존의 협업 필터링과 비교한 결과 전체적으로 추천의 성과가 향상되어 향후 전시 환경에서의 부스 추천시스템에 적용하여 관람객의 니즈를 충족시킬 것으로 기대된다.

.....

논문접수일 : 2010년 11월 23일

게재확정일 : 2010년 12월 07일

교신저자 : 김혜경

1. 서 론

전시회는 전시업체가 새로운 상품이나 서비스를 관람객에게 알리기 위해 개최되는 것으로 효과적인 마케팅 수단으로 중요한 역할을 수행한다(John et al., 2001; Bellizi and Lipps, 1984). 이러한 전시회는 전시 후 구전 효과가 크기 때문에 관람객들의 다양한 니즈를 충족시키는 것이 필요하다(이희찬과 한진영, 2005).

최근의 다양한 관람객의 니즈를 충족시키기 위

해 다양한 유비쿼터스 기술이 전시회에 응용되고 있다. 특히, 블루투스, RFID 등의 무선 센싱 네트워크 기술로는 기존의 전시회에서는 파악하기 어려웠던 관람객의 동선 등과 같은 관람객의 정보를 얻을 수 있다(전정호와 이경진, 2009). 이러한 정보는 개인의 선호와 방문 패턴을 발견하는 데 도움이 되며 이를 통해 각 개인에게 적합한 정보를 선별 및 제공함으로써 관람객의 다양한 욕구를 만족시킬 수 있다. 하지만 현재 전시회에 적용되고 있는 유비쿼터스 기술은 단순히 부스에 관련된 정보를 전

* 본 논문은 지식경제부 지식서비스 유비쿼터스 센싱 네트워크(USN) 산업원천기술개발사업으로 지원된 연구임 (10035426).

달하는 역할만 수행하고 있다. 특히, 전시회의 대규모화로 인하여(예를 들어, IT분야에서 대표적인 CeBIT 전시회의 경우 2010년에 68개국에서 4,157개의 업체들이 참여하였다) 관람객이 자신의 선호에 적합한 부스를 선택하는 데 있어 복잡성이 있음에도 불구하고 단지 관람객이 사전에 요청한 정보만을 제공한다. 즉, 현재 전시회에서 이루어지고 있는 서비스는 개별 관람객의 선호가 반영되지 않는 단순 정보 제공에 불과하여 관람객의 니즈를 충족시키기에는 한계가 있다. 이러한 문제를 해결하기 위한 방법으로 추천 시스템이 대두되고 있다(이연남과 권오병, 2008; Herlocker et al., 2004).

추천시스템은 개인화된 정보 탐색 및 개인의 의사결정을 돕는 도구로 방대한 정보원천에서 특정 개인의 선호에 맞는 정보를 선택하여 제공한다(김재경과 안도현과 조윤호, 2003; Sarwar et al., 2000). 따라서 추천시스템은 전시회에서 개인의 선호를 추론하고 개인의 선호에 부합하는 부스를 추천하기 때문에 관람객의 니즈를 충족시킬 수 있다. 이러한 추천 시스템 중 협업 필터링(Collaborative Filtering, CF)은 아마존, 넷플릭스와 같은 실제 적용에서도 가장 성공적인 방법론으로 평가받고 있다(Amer-Yahia, Lakshmanan, and Yu, 2009).

하지만 전통적인 협업 필터링은 추천 과정에서 사용자의 선호 변화를 고려하지 않고 사용자의 평가 정보에 의존하여 추천함에 따라 전시 환경에 적용하기에 한계가 있다. 따라서 본 연구에서는 전시 환경에 적합하도록 순차 연관 규칙(Sequential Association Rule, SAR) 기법을 이용하여 관람객의 부스 방문 순서를 고려한 방법론을 제안한다. 즉, 관람객의 부스 방문기록에 순차 연관 규칙 기법을 이용하여 관람객의 부스 방문 기록 내에 존재하는 규칙의 신뢰도를 계산한 후 이를 바탕으로 목표 관람객의 선호에 부합하는 부스를 추천하는

방법론을 제안한다.

본 연구는 다음과 같이 구성된다. 다음 장에서는 관련 연구를 제시하고 제 3장에서는 본 연구가 제안하는 추천 방법론을 설명하며, 제 4장에서는 예제를 통해 방법론을 보충 설명한다. 제 5장에서는 추천 방법론을 검증하기 위한 실험 설계와 그 결과를 보여주고 마지막 장에서는 본 연구의 결론과 한계점에 대하여 논의한다.

2. 관련 연구

2.1 협업 필터링

협업 필터링(Collaborative Filtering)은 고객들의 상품에 대한 평가를 이용하여 일상생활에서 가족, 친구, 동료들의 경험을 통한 구전 효과를 자동화 한 것으로 선호도가 유사한 고객 그룹을 결정하며 그들의 상품 구매 기록을 바탕으로 추천되는 상품을 결정한다(김재경과 김혜경과 최일영, 2009; Sarwar et al., 2000). 하지만 기존의 협업 필터링은 전시 환경에 적용하기에 한계가 있다.

전시회가 열리는 장소는 입구 및 출구가 정해져 있으며 관람 편의 및 전시 주관 업체의 목적에 따라 몇 가지의 경로가 존재한다. 또한, 전시회를 방문한 관람객은 자신의 성향에 따라 진행 방향을 선택한다(Amer-Yahia, Lakshmanan, and Yu, 2009). 따라서 관람객의 부스 방문 여부 뿐만 아니라 부스를 방문한 순서에도 관람객의 선호가 반영되지만 기존의 협업 필터링의 경우에는 구매 여부 또는 방문 여부의 정보만 반영될 뿐 순서는 전혀 고려되지 않는다(Lu, Agarwal, and Dhillon, 2009).

또한, 전시 환경에서 전시업체는 제한된 기간 동안 홍보 목적을 달성하기 위해서 관람객과 끊임없이 상호작용을 하게 되고 상호작용의 결과에 따라

관람객의 선호가 변하는 현상이 발생한다(Sandra, 2002). 상호작용에 따라 실시간으로 변화되는 관람객의 선호 뿐만 아니라 전시회를 방문한 관람객은 신제품과 신기술의 수준을 파악이라는 목적을 가지고 있어, 대개 이전에 들렀던 부스 방향 보다 새로운 방향을 선택하는 경향이 있다(Blythe, 1999). 하지만 기존의 협업 필터링은 목표관람객이 추천을 요청할 때 이전의 기록 전체를 바탕으로 추천 서비스를 제공함에 따라 전시회 내에서 변화되는 관람객의 선호가 적극적으로 반영되지 못한다는 한계가 있다. 즉, 협업 필터링은 목표 관람객의 부스 방문 기록 전체와 유사한 관람객을 추천 절차에 사용함에 따라 관람객의 선호가 변화되어도 이전의 기록에 의존하여 추천 서비스를 제공한다. 또한, 목표 관람객이 이전에 방문하지 않았던 부스 전체를 추천의 대상으로 추천 절차를 진행함에 따라 목표 관람객이 이전에 방문했던 방향으로 추천이 제공될 가능성이 존재하여 새로운 정보의 탐색이라는 관람객의 목적에 부합되지 않아 추천의 정확도가 떨어진다.

2.2 순차 연관 규칙

고객의 선호 변화를 반영하기 위해 다양한 기법이 추천 시스템에 응용되고 있다. 기존의 추천 방법론에서 순서를 고려하기 위해 대표적으로 마코브 모델(Markov Model), 연관성 규칙(Association Rule), 순차 연관 규칙(Sequential Association Rule, SAR) 등이 쓰이고 있다.

마코브 모델은 전이 확률을 통해 순서 의존도를 밝혀내는 간단한 모델이다. 초기 마코브 모델은 미래의 상태가 현재 시점의 영향을 받는다는 가정하에 미래 시점의 행동을 예측하는 모델로 바로 직전 시점에 대한 고려만을 한다는 점에서 한계를

가졌다. 따라서 이러한 한계점을 해결하기 위해 여러 단계에 걸친 과거 시점까지 고려할 수 있는 모델들이 연구되고 있다. 하지만 모델 자체의 복잡성으로 인하여 과거 시점의 수가 증가함에 따라 계산량이 기하급수적으로 증가한다. 또한, 모델을 위해 다른 기법에 비해 많은 양의 데이터가 필요하여 데이터의 양이 충분치 않은 영역에 적용하기에는 한계가 있다(Kim et al., 2004). 일반적으로 전시회는 일시적으로 열리며 시간에 민감한 시장임에 따라 모델을 구축하기 위해 사용될 수 있는 데이터의 양이 충분하지 않기 때문에 마코브 모델을 이용하여 부스 추천 시스템에 이용하는데 한계가 있다.

연관성 규칙은 선행 항목집합과 후행항목집합의 동시 구매 정보에 기반한 기법이다(Mobasher et al., 2000). 연관성 규칙에서 가장 대표적으로 사용되는 기법은 Apriori 알고리즘으로 동시에 빈번하게 발생하는 항목집합, 즉 신뢰도가 높은 규칙을 찾아내어 추천에 사용한다(Agrawal et al., 1993). 연관성 규칙은 마코브 모델에 비해 정확도는 떨어지지만 높은 재현율을 가져 설명력이 높은 편이며, 최소 지지도 이상의 빈발 항목집합만을 대상으로 규칙을 생성함에 따라 계산량이 감소된다. 하지만 연관성 규칙의 경우 단순히 동시 구매 정보만을 고려하기 때문에 고객의 선호 변화는 반영할 수 없는 문제점이 있다.

순서를 고려한 연관성 규칙인 순차연관 규칙의 기본적인 가정은 선행 항목집합이 후행 항목집합에 시간적으로 우선해야 한다. 최근의 순차 연관 규칙에 관한 연구들은 제약 조건을 사용하여 추천의 성과를 높이고 있다. 예를 들어, Srikant et al.(1997)은 매번 거래의 순서를 고려하는 것이 아니라 일정 시간 동안의 거래들을 항목집합으로 구성하여 항목 집합간의 순서를 통해 규칙을 생성하

며 최대 시간 간격과 최소 시간 간격을 설정하여 모든 시간을 대상으로 거래의 순서를 고려하는 것과 유사한 성과를 내는 방법론을 제안하였다. Ng et al.(1998)은 순서를 가지는 데이터 중에서 풍부한 양의 정보를 가지는 일정 시간 동안의 거래만을 사용하여 빈번하게 발생하는 규칙들을 추출해내는 방법론을 제안하였으며 또한 Zaki(2000)는 고객의 거래 데이터 특성에 따라 다르게 시간 간격을 부여하여 순차 연관 규칙의 정확도를 높이는 방법론을 제안하였다. 이러한 방법론들은 시간 간격을 제약조건으로 하여 빈번하게 발생한 순서 규칙을 발견한다. 그러나 기존의 순차 연관 규칙은 추천 과정에서 전체 규칙을 모두 고려해야 함에 따라 전시회의 규모가 대형화되면서 계산량의 증가라는 한계를 지닌다.

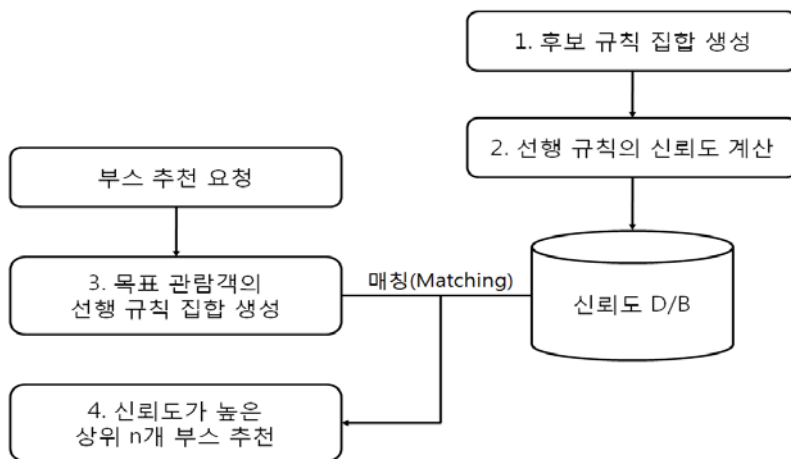
본 연구에서는 이러한 문제점을 해결하기 위해 Harms et al.(2002)의 연구를 기반으로 하여 순차 연관 규칙을 이용한 추천 방법론을 제시한다. Harms et al.(2002)은 빈번하게 발생하는 선행 사건만을 대상으로 신뢰도를 계산하는 것이 전체 사건에서 신뢰도를 계산하는 것보다 더 효과적이라는 가정

하에 순차 연관 규칙을 생성했다. 따라서 본 연구에서는 위와 같은 가정을 바탕으로 후행 사건을 부스로 정의하고 후행 사건에 시간적으로 선행하는 선행 규칙 중 일정 수준 이상 출현하는 선행 규칙만을 대상으로 후보 선행규칙집합을 구성함에 따라 대형화되는 전시 환경에 적합한 방법론을 제시한다. 또한, 후보 선행 규칙 집합 및 후행 사건에 대한 신뢰도를 미리 신뢰도 데이터 베이스에 저장한 뒤 실제 추천에서는 목표 관람객의 부스 방문 기록을 신뢰도 데이터 베이스와 매칭하여 추천 정보를 제공함으로써 유비쿼터스 환경에 적합한 추천 방법론을 제시한다.

3. 순차 연관 규칙을 이용한 추천 방법론

3.1 개요

관람객의 부스 방문 순서는 관람객의 선호가 반영 된다. 따라서 본 연구에서는 관람객의 방문 순서를 고려한 추천 방법론을 제시하기 위해 순차 연관 규칙을 이용한다. 본 연구에서 제시하는 방법론



<그림 1> 프레임워크

의 프레임 워크(Framework)는 <그림 1>과 같다.

3.2 신뢰도 데이터 베이스의 구성

신뢰도 데이터 베이스의 구성은 2단계로 구성된다. 1단계는 후보 규칙 집합 생성으로 관람객의 부스 방문 기록에서 보편적으로 도출되는 선행 규칙을 후보 규칙 집합으로 구성한다. 다음 단계는 후보 규칙 집합의 선행 규칙이 후행 사건을 포함할 확률을 계산하여 추천에 사용될 신뢰도 데이터 베이스를 구성한다.

• 단계 1 : 후보 규칙 집합 생성

본 연구에서 제시하는 방법론의 적용을 위하여 관람객의 부스 방문 기록($u_i(S)$)은 다음과 같이 정의한다.

$$u_i(S) = \{a_1, a_2, a_3, \dots, a_n\} \quad (1)$$

여기에서 a_n 은 관람객이 방문한 부스를 의미하며 n 은 관람객이 방문한 총 부스의 수를 의미한다.

따라서 $u_i(s)$ 는 관람객 u_i 의 부스 방문 순서 및 방문한 부스의 수에 따라 관람객 별로 정의된다. 예를 들어, 관람객 u_1 이 $b_1 \rightarrow b_5 \rightarrow b_3 \rightarrow b_2$ 의 순으로 부스를 방문하였다면, $n = 4$ 로 $a_1 = b_1, a_2 = b_5, a_3 = b_3, a_4 = b_2$ 가 되어 관람객 u_1 의 부스 방문 기록 $u_1(s) = \{b_1, b_5, b_3, b_2\}$ 이 된다.

순차 연관 규칙은 선행 규칙 집합이 시간적으로 후행 사건에 선행해야 한다는 가정하에 선행 규칙과 후행 사건간의 빈도와 신뢰도를 고려한다. 본 연구에서는 전시 환경에 적용하기 위해 관람객이 추천을 요청한 시점에 위치하고 있는 부스 이전의 방문 기록을 바탕으로 빈도와 신뢰도를 도출하여

다음으로 방문할 부스를 추천한다. 먼저 선행 규칙(R_m^p)의 정의는 다음과 같다.

$$R_m^p \subset v_{all}(S) \quad (p = 1, 2, 3, \dots, \max(n) - 1) \quad (2)$$

p 는 m 번째 선행 규칙이 포함하는 부스의 수를 의미하며 p 의 최대값은 후행 사건에 대한 신뢰도 계산을 위해 모든 관람객별 부스 방문 수 중 최대값 즉, $\max(n)$ 에서 1을 뺀 값으로 정의한다. 하지만 모든 관람객의 모든 규칙을 선행 규칙으로 고려할 경우 고려해야 할 규칙의 수가 기하급수적으로 늘게 되어 모델 구축의 계산량을 증가시킬 뿐만 아니라 관람객들의 부스 방문 규칙에서 거의 존재하지 않는 규칙까지 고려하게 되어 추천의 성과가 저하된다. 따라서 후행 사건에 선행하는 규칙 중 일정 수준 이상의 빈도(frequency)를 가지는 선행 규칙 즉, 보편적으로 도출되는 규칙들을 대상으로 추천이 이루어질 수 있도록 선행 규칙을 제한해야 한다. 본 연구에서는 빈도($fr(R_m^p \Rightarrow b)$)를 다음과 같이 정의한다.

$$fr(R_m^p \Rightarrow b) = \frac{n(R_m^p \Rightarrow b)}{n(b)} \quad (3)$$

여기에서 $n(b)$ 는 후행 사건으로 부스 b 를 포함하는 모든 선행 규칙의 수를 의미하며 $fr(R_m^p \Rightarrow b)$ 은 후행사건 b 를 포함하는 선행 규칙 R_m^p 의 수이다. 즉, 빈도의 정의는 후행 사건 b 를 포함하는 모든 선행 규칙들 중 해당 선행 규칙 R_m^p 의 비율을 의미한다.

위와 같은 빈도를 바탕으로 일정 수준(a) 이상의 빈도를 가지는 선행 규칙만을 대상으로 후보 규칙 집합을 구성한다. 후보 규칙 집합(Cr)의 정의

는 다음과 같다.

$$\begin{cases} fr(R_m^p = \max(fr(R_m^p \Rightarrow b)) \\ Cr = \{R_m^p | fr(R_m^p) \geq \alpha\} \end{cases} \quad (4)$$

선행 규칙의 빈도값 $fr(R_m^p)$ 은 후행 사건을 포함하는 선행 규칙의 빈도 값 중 최대값으로 정의되며 이 빈도값이 일정 수준 이상인 선행 규칙으로 후보 규칙 집합을 구성하게 된다.

• 단계 2 : 선행 규칙의 신뢰도 계산

일정 수준 이상의 빈도를 가지는 선행 규칙을 후보 규칙 집합으로 구성한 후 각 규칙들의 신뢰도(confidence)를 계산하여 향후 목표 고객의 추천 절차에 사용하게 된다. 신뢰도($c(R_m^p \Rightarrow b)$)는 다음과 같이 정의된다.

$$c(R_m^p \Rightarrow b) = \frac{n(R_m^p, b)}{n(R_m^p)}, (R_m^p \in Cr) \quad (5)$$

$n(R_m^p)$ 은 후보 규칙 집합 내에서 선행 규칙의 수를 의미하며 $n(R_m^p, b)$ 은 선행 규칙과 후행 사건, 즉 부스 b 가 동시에 나타나는 수를 의미한다. 따라서 신뢰도는 선행 규칙 R_m^p 에 대한 후행 사건의 조건부 확률을 말한다.

위에서 제시한 두 단계를 거친 데이터는 신뢰도 데이터 베이스에 저장하여 관람객이 추천 정보를 요청할 때에 이 신뢰도 데이터 베이스를 탐색하여 추천 정보를 제공한다.

3.3 부스의 추천

부스 추천 절차는 다음과 같이 2단계로 구성되

어 있다. 방문한 관람객이 추천 정보를 요청하게 되면 3단계로 해당 관람객의 선행 규칙을 생성한다. 해당 관람객의 선행 규칙이 생성되면 4단계로 신뢰도 데이터 베이스 내에 규칙이 존재하는지를 탐색하여 존재하는 규칙의 신뢰도가 높은 상위 n 개의 부스를 추천한다.

• 단계 3 : 목표 관람객의 선행 규칙 집합 생성

유비쿼터스 전시 환경에서 관람객이 특정 부스에 위치한 상태에서 휴대 단말기로 다음으로 방문할 부스에 대한 추천을 요청하게 되면 해당 관람객이 이전에 방문했던 부스 방문 기록을 바탕으로 관람객의 선행 규칙을 구성한다. 위에 제시한 부스 방문 기록의 정의와 같이 목표 관람객 u_t 의 방문 기록을 $u_t(S) = \{a_1, a_2, a_3, \dots, a_n\}$ 이라 정의하면 목표 관람객의 선행 규칙(r_m^q)은 다음과 같이 생성된다.

$$r_m^q \subset u_t(S) (q=1, 2, 3, \dots, n) \quad (6)$$

여기에서 q 는 m 번째 선행 규칙이 포함하는 부스의 수를 의미하며, q 의 최대값은 목표 관람객이 추천을 요청한 시점에 위치한 부스와 그 이전에 방문했던 부스의 수 n 이다.

• 단계 4 : 신뢰도가 높은 상위 n 개 부스 추천

목표 관람객의 선행 규칙이 생성된 후에는 먼저 신뢰도 데이터 베이스 안에서 각 선행 규칙이 존재하는지 여부를 탐색한다. 만약, 목표 관람객의 선행 규칙이 신뢰도 데이터 베이스 안에 존재하지 않게 되면 이는 보편적인 규칙이 아닌 것으로 판단하여 추천에 사용하지 않는다.

추천에 사용될 목표 관람객의 선행 규칙이 결정 되면 목표 관람객이 아직 방문하지 않은 부스만을 후행 사건 대상으로 하여 신뢰도 데이터 베이스 안에서 목표 관람객의 각 선행 규칙 r_m^a 이 가지는 신뢰도를 탐색한다. 이에 따라, 각 선행 규칙이 후행 사건에 가지는 신뢰도 중 높은 순으로 상위 n 개의 부스를 추천한다.

4. 예제

이번 장에서는 본 연구에서 제시한 방법론의 이해를 돕기 위한 간략한 예제를 제시한다. 예제에 사용되는 데이터는 <표 1>과 같다.

<표 1> 데이터 예제

ID	#1	#2	#3	#4
u_1	B1	B2	B3	B4
u_2	B5	B2	B1	B3
u_3	B1	B2	B4	B5
u_4	B1	B5	B2	B4

데이터 예제는 총 5개의 부스를 가지고 있으며 4명의 관람객이 각각 4개의 부스를 방문하였다고 가정한다. 또한 #number는 관람객의 부스 방문 순서를 의미한다.

4.1 신뢰도 데이터 베이스의 구성

단계 1 : 후보 규칙 집합 생성

<표 1>과 같은 부스 방문 기록이 있다고 가정하였을 때, 이를 바탕으로 후보 규칙 집합을 구성한다. 본 예제에서는 총 5개의 부스에 대한 관람객의 방문 기록이 있으며 관람객 중 가장 많은 부스

를 방문한 관람객의 부스 방문 수는 4로 p 의 최대 값은 3으로 선행 규칙 R_m^p 은 총 25개의 규칙이 생성된다.

선행 규칙을 정의한 후 후보 규칙 집합의 구성을 위해 빈도를 계산한다. 본 연구에서 정의하는 빈도는 후행 사건을 포함하는 선행 규칙의 수임에 따라 빈도는 총 70번 계산한다. 빈도를 모두 계산한 후 각 선행 규칙별로 가장 높은 빈도를 가지는 규칙을 선정하여 해당 선행 규칙의 빈도값으로 사용한다. <표 2>는 후행 사건을 포함하는 빈도의 계산 후 각 선행 규칙별 최대값을 내림차순으로 정렬한 것 중 일부이다.

<표 2>에 제시된 선행 규칙 중 일정 수준(a)를 만족하는 선행 규칙으로 후보 규칙 집합을 생성한다. 본 예제에서는 $a = 0.1$ 로 정하여 후보 규칙 집합(C_r)을 구성한다.

<표 2> 빈도의 계산

선행 규칙(\Rightarrow 후행 사건)	빈도
B5(\Rightarrow B2)	0.5(2/4)
B1(\Rightarrow B3)	0.27(3/11)
B2(\Rightarrow B1)	0.25(1/4)
B3(\Rightarrow B1)	0.25(1/4)
B1, B5(\Rightarrow B2)	0.25(1/4)
B2, B5(\Rightarrow B1)	0.25(1/4)
B1, B2(\Rightarrow B3)	0.18(2/11)
B1, B4(\Rightarrow B5)	0.125(1/8)
B2, B4(\Rightarrow B5)	0.125(1/8)
B1, B2, B4(\Rightarrow B5)	0.125(1/8)
B1, B2, B5(\Rightarrow B3)	0.09(1/11)
...	...

단계 2 : 선행 규칙의 신뢰도 계산

후보 규칙이 생성되면 다음으로 후보 집합에 속

하는 선행 규칙 뒤에 빈번하게 방문하는 부스를 발견하기 위하여 선행 규칙 뒤에 후행 사건이 발생할 신뢰도를 계산한다.

단계 1에서 α 를 0.1로 정함에 따라 후보 규칙 집합은 총 10개의 선행 규칙으로 구성되었다. 이를 바탕으로 선행 규칙에 대한 각 부스별 신뢰도를 계산하게 된다. <표 3>은 신뢰도 계산 결과 중 각 선행 규칙에서 가장 높은 신뢰도를 가지는 후행 사건을 나열하였다.

<표 3> 신뢰도의 계산

선행 규칙	후행 사건	신뢰도
B5	B2	0.66(2/3)
B1	B2	0.75(3/4)
	B4	0.75(3/4)
B2	B4	0.75(3/4)
B3	B4	0.5(1/2)
B1, B5	B2	0.5(1/2)
	B3	0.5(1/2)
	B4	0.5(1/2)
B2, B5	B1	0.5(1/2)
	B3	0.5(1/2)
	B4	0.5(1/2)
B1, B2	B4	0.75(1/2)
B1, B4	B5	1(1/1)
B2, B4	B5	1(1/1)
B1, B2, B4	B5	1(1/1)

4.2 부스의 추천

단계 3 : 목표 관람객의 선행 규칙 집합 생성

본 연구에서 제시하는 방법론에서는 추천 절차를 진행하기 위해 먼저 목표 관람객의 부스 방문 기록을 바탕으로 하여 선행 규칙 집합을 생성한다. 목표 관람객 u_t 의 부스 방문 기록이 $B3 \rightarrow B2 \rightarrow B4$ 와 같은 경우 부스 방문 기록은 $u_t(S) = \{B3,$

$B2, B4\}$ 로 정의된다. 본 예제에서 목표 관람객의 부스 방문 수가 총 3개로 부스 방문 기록에서는 7개의 선행 규칙이 다음과 같이 도출 된다.

$$r_1^1 = \{B3\}$$

$$r_2^1 = \{B2\}$$

$$r_3^1 = \{B4\}$$

$$r_4^2 = \{B3, B2\}$$

$$r_5^2 = \{B3, B4\}$$

$$r_6^2 = \{B2, B4\}$$

$$r_7^3 = \{B3, B2, B4\}$$

단계 4 : 신뢰도가 높은 상위 n개 부스 추천

목표 관람객의 선행 규칙을 생성한 후에는 신뢰도 데이터 베이스 내에 해당 선행 규칙이 있는지 탐색한다. 본 예제에서의 신뢰도 데이터 베이스 내에 10개의 후보 선행 규칙이 존재하며 목표 관람객의 각 선행 규칙이 후보 규칙 집합내의 선행 규칙과 일치하게 되면 추천에 사용한다. 따라서 후보 선행 규칙 내에 존재하지 않는 목표 관람객의 선행 규칙 $r_3^1, r_4^2, r_5^2, r_7^3$ 은 추천에 사용되지 않는다.

추천 절차의 마지막 과정으로 신뢰도 데이터 베이스에 존재하는 목표 관람객의 선행 규칙들 r_1^1, r_2^1, r_6^2 뒤에 방문한 가능성이 높은 부스, 즉 신뢰도가 높은 후행 사건을 목표 관람객에게 추천한다. 단, 목표 관람객이 이전에 방문했던 부스는 추천 대상에서 제외한다. 각 선행 규칙별로 신뢰도 데이터 베이스 내에 존재하는 후행규칙에 대한 신뢰도는 <표 4>와 같다.

따라서 본 예제에서는 목표 관람객에게 신뢰도가 가장 높은 B5를 추천한다.

다음 장에서는 실제 전시회에 방문한 관람객들의 데이터를 이용하여 본 연구에서 제시하는 방법

론의 성과를 전통적인 협업 필터링과 비교하여 측정한다.

<표 4> 목표 관람객의 선행 규칙별 신뢰도

선행 규칙	후행 사건	신뢰도
B3	B1	0(0/4)
	B5	0(0/4)
B2	B1	0.25(1/4)
	B5	0.25(1/4)
B2, B4	B1	0(0/1)
	B5	1(1/1)

로 데이터를 한정한다. 또한, 부스 방문 수가 너무 많은 관람객의 경우 특정한 선호를 가지지 않고 대부분의 부스를 방문하였다고 가정하여 방문 수가 총 부스의 75%이상, 즉 18개 이상의 부스를 방문한 고객의 데이터는 제외하였다. 최종적으로 전통적인 협업 필터링과 SAR 기법을 이용한 방법론의 성과 비교를 위해 사용된 데이터를 요약하면 다음과 같다.

<표 5> 데이터 특성

관람객 수	부스 수	관람객 별 부스 방문 수
316	24	7~17

5. 실험

5.1 데이터 특성

본 연구에서 제시한 방법론의 성과를 측정하기 위해 사용하는 데이터는 'L'사가 주관한 2008년 IT 관련 전시회의 부스 방문 기록으로 RFID 기술을 활용하여 사전 등록 또는 현장등록한 관람객을 대상으로 RFID칩이 내장된 등록 카드를 발부하여 관람객이 부스 방문 시에 부스에 마련된 RFID 리더에 태깅(Tagging)을 유도하여 데이터를 수집하였다.

부스 방문 기록 원본 데이터는 총 24개의 부스와 822명의 관람객으로 구성되어 있으며, 각 관람객별 부스 방문 수는 1~22개로 방문 순서에 대한 정보를 포함하고 있다. 본 실험에서는 추천 성과의 비교 뿐만 아니라 부스 추천 수와 추천에 사용되는 데이터 수에 따른 성과의 변화 패턴도 살펴본다. 그러나 기존의 전통적인 협업 필터링과 본 연구에서 제시한 방법론 간의 성과 비교를 위해 원본 데이터를 그대로 사용하는 것은 한계가 있다. 따라서 최소 7개 이상의 부스를 방문한 관람객으

5.2 실험 방법

일반적으로 성과 측정을 위해 표본 데이터를 학습용 집합(training set)과 평가용 집합(test set)으로 분할한다. 이는 기법의 구축과 성과 측정에 동일한 데이터를 사용할 경우 편향(bias)이 발생할 가능성이 높기 때문이다. 따라서 본 연구에서는 관람객 316명의 데이터를 무작위로 분할하여 219명(69.1%)의 부스 방문 기록은 기법을 구축하는 데 사용하고 97명(30.9%)의 부스 방문 기록은 성과 측정을 위한 평가용 집합으로 사용하였다. 그리고 순차 연관 규칙은 일정 수준 이상의 후행 사건을 포함하는 선행 규칙만을 대상으로 추천을 제공하기 때문에 본 실험에서는 선행 규칙의 빈도 수준 기준값(α)을 0.2로 설정하여 순차 연관 규칙을 생성하였다.

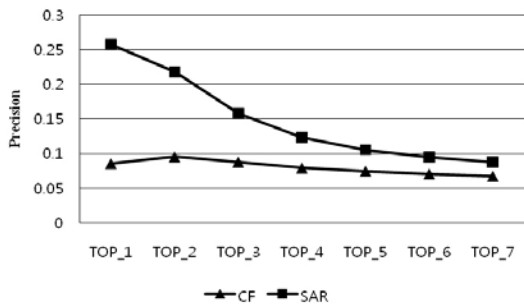
추천 시스템의 성과를 평가하기 위한 지표로는 대표적으로 정확도(Precision)와 재현율(Recall)이 사용된다(Billsus and Pazzani, 1998; John et al., 2001). 이 연구에서는 이전 부스 방문 정보를 바탕으로 부스를 추천하면, 추천한 부스중에 실제 고객

이 방문한(1개의) 부스가 포함되는지를 평가하므로 정확도(precision) 만이 평가지표로 사용된다. 추천한 부스와 관람객이 방문한 모든 부스를 대상으로 평가한다면, 정확도와 재현율이 다 사용되어야 하나, 전시장에서의 부스 관람은 순서가 중요하므로 다음 방문할 부스만을 평가하므로 정확도만을 평가 지표로 사용한다.

5.3 실험 결과 및 분석

5.3.1 추천 부스 수에 따른 성과 비교

전통적인 협업필터링은 구매가능점수(*Purchase Likelihood Score, PLS*)에 따라 상위 n 개의 부스를 추천하며(김재경과 김혜경과 최일영, 2009; Sarwar et al., 2000) 순차 연관 규칙은 선행 규칙에 대한 후행 사건의 신뢰도를 기준으로 상위 n 개의 부스를 추천한다. 따라서 추천 시스템의 성과 비교를 위해 추천 부스 수에 따른 정확도의 변화를 살펴 보았다. <그림 2>는 추천 부스 수에 따른 각 추천 시스템의 정확도를 나타내는 그래프로 전체적으로 순차 연관 규칙을 적용한 방법론이 전통적인 협업 필터링 보다 높은 성과를 보인다.



<그림 2> 추천 부스 수에 따른 정확도(Precision)

추천 부스의 수에 따른 순차 연관 규칙과 기존의 협업필터링의 정확도 차이는 다음과 같다.

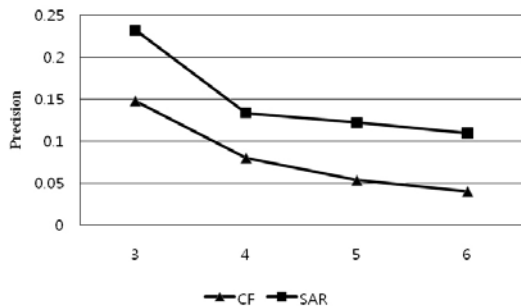
첫째, 부스 추천 수에 따른 변화 패턴에서 정확도가 감소하는 패턴의 차이가 있다. 본 연구에서 제시하는 방법론의 경우 추천 부스 수가 증가함에 따라 추천 성과가 지속적으로 감소하는 패턴을 보이는 반면 협업 필터링은 상위 2개를 추천하는 경우의 정확도가 1개를 추천하는 경우보다 높다. 순차 연관 규칙의 지속적인 정확도 감소는 추천 과정에서 사용하는 데이터의 특성에서 기인한다. 협업 필터링은 선호가 유사한 관람객들의 부스 방문 기록을 바탕으로 추천 서비스를 제공하는 반면에 순차 연관 규칙은 전체 관람객들의 부스 방문 기록을 바탕으로 빈번하게 발생하는 규칙을 추천한다. 따라서 순차 연관 규칙을 이용하는 경우 전체 관람객들의 관람 패턴과 다른 경향을 보이는 관람객의 선호 추론에 한계를 지니 부스 추천 수를 늘리게 되더라도 추천이 적중하지 못하게 된다. 반면에, 협업 필터링은 추천 과정에서 선호가 유사한 관람객들의 정보를 사용함에 따라 전체 관람객과 다른 패턴을 지닌 관람객에 대한 추천 정보 제공이 가능하지만 방문 순서가 고려되지 않아 현재 시점에서 바로 다음 시점에 나올 가능성이 낮은 부스를 추천할 수 있다. 따라서 협업 필터링은 추천하는 부스를 증가시킴에 따라 다음 시점에 나올 가능성이 높은 부스가 추천 정보에 포함될 확률이 높아진다. 협업 필터링의 경우 추천하는 부스의 증가에 따른 정확도의 감소를 상쇄시켜 상위 1개를 추천할 때의 정확도와 7개를 추천할 때의 정확도 차이가 크지 않다.

둘째, 순차 연관 규칙이 가지는 특징은 정확도의 급속한 감소 패턴에서도 보인다. 그 원인은 앞서 제시한 전체 관람객과 다른 경향을 보이는 관람객에 대한 추천 적중 실패 뿐만 아니라 상위 1개를 추천하였을 경우의 높은 정확도도 영향을 끼친다. 상위 1개의 추천이 적중한 관람객에게 상위 2

개의 부스를 추천하였을 경우에는 정확도가 100%에서 50%로 감소한다. 마찬가지로 상위 3개의 부스를 추천하게 되면 33.3%의 정확도를 가지게 되면서 부스 추천 수가 증가함에 따라 정확도가 지수함수적으로 감소한다. 따라서 그림 2에서 관찰되는 순차 연관 규칙의 변화 패턴이 일정한 감소 패턴보다는 지수함수적인 감소 패턴을 보이는 것은 상위 1개를 추천하였을 때 적중한 고객의 비율, 즉 높은 정확도의 영향이 크다.

5.2.2 부스 방문 누적에 따른 성과의 변화

협업필터링과 순차 연관 규칙의 성과를 비교하기 위한 다른 방법으로 목표 관람객의 부스 방문 수 증가에 따른 성과의 변화를 측정하였다. 본 실험에서는 추천 시스템에서 사용되는 목표 관람객의 부스 방문 수의 증가에 따른 성과의 변화를 측정하기 위해 목표 관람객의 부스 방문 수가 3개일 때부터 6개가 될 때까지의 정확도를 측정하였다. 예를 들어, 목표 관람객의 부스 방문 수가 3개일 때는 바로 다음 부스 방문인 4번째 부스 방문의 추천이 적중한지의 여부를 측정한다. <그림 3>은 부스 방문 수의 증가에 따른 정확도의 변화 그래프를 보여준다.



<그림 3> 부스 방문 수에 따른 정확도(Precision)

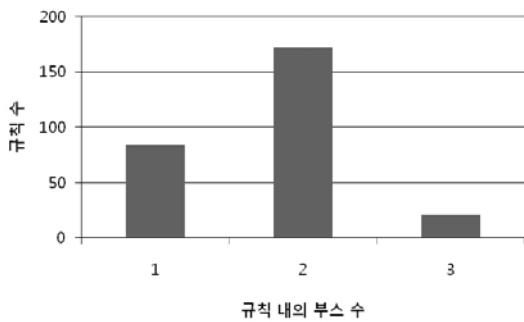
추천 부스의 증가에 따른 정확도 비교와 마찬가지로 본 연구에서 제시한 방법론이 협업 필터링보다 전체적으로 높은 정확도를 보여준다.

부스 방문 수에 따른 정확도 비교에서 관찰되는 특징은 협업 필터링과 순차 연관 규칙 모두 목표 관람객의 부스 방문 누적에 따라 정확도가 현저히 저하된다는 것이다. 추천 시스템에서 목표 관람객의 부스 방문 기록은 관람객의 선호를 추론하기 위해 사용된다. 따라서 목표 관람객의 부스 방문 기록 수의 증가는 선호를 추론하기 위해 사용되는 데이터의 양이 늘어나는 것을 의미한다. 일반적으로 추천시스템의 성과는 목표 관람객의 데이터가 늘어남에 따라 증가하는 경향을 보이지만 일정 시점 이후에 감소된다. 하지만 <그림 3>에서 두 추천 방법 모두 추론에 사용되는 데이터의 양이 증가함에 따라 정확도가 감소되는 원인은 다음과 같다.

관람객이 전시회를 방문하면 기존의 시장에서 볼 수 없었던 새로운 상품이나 서비스를 경험하게 된다. 즉, 전시회에 참여하는 업체는 기존에 경험할 수 없었던 새로운 상품이나 서비스를 제공하면서 관람객의 선호는 부스를 방문할 때마다 변화할 가능성이 높다. 따라서 관람객이 전시회의 입장할 때부터의 모든 데이터를 추천시스템에서 관람객의 선호를 추론하는데 사용하는 것은 오히려 추천시스템의 성과를 저하시키는 원인이 된다. 전시 환경에서 추천시스템의 성과를 높이기 위해서는 목표 관람객이 방문한 전체 부스의 데이터를 사용해 추천을 제공하는 것보다 목표 관람객이 추천을 요청하기 이전의 일정 시점으로 제한하여 추천을 제공하는 것이 더 효과적이다. 이러한 전시환경의 특징은 <그림 4>에서도 나타난다. <그림 4>는 순차 연관 규칙에서 사용된 선행 규칙의 부스 수를 나타낸 그래프이다.

순차 연관 규칙에서는 일정 수준 이상의 빈도를

가지는 선행 규칙만을 대상으로 추천 절차를 진행한다. 본 실험에서는 선행 규칙의 빈도가 0.2(a) 이상인 선행 규칙만을 대상으로 추천 절차를 진행하였다. 순차 연관 규칙을 이용해 생성된 규칙은 총 277개로 추천 절차에 주로 사용된 선행 규칙은 부스를 1~2개(92.4%)를 포함하는 규칙이다. 또한, 선행 규칙이 4개 이상인 규칙은 추천 절차에 사용되지 않았다. 즉, 추천 절차에 사용되는 목표 관람객의 데이터가 1~3개만 되더라도 가장 높은 신뢰도를 도출할 수 있다.



<그림 4> 선행 규칙 내의 버스 수

6. 결론 및 한계점

전시회의 성공을 위해서는 전시회 관람의 주체가 되는 관람객의 니즈를 충족시켜 만족을 제공해야 한다. 최근 유비쿼터스 기술의 발전으로 기존에 획득할 수 없었던 정보의 획득이 가능하여 고객의 선호에 부합하는 개인화 서비스를 제공할 수 있는 기반이 마련되었지만 유비쿼터스 환경이 구축된 전시회의 서비스는 개인의 선호를 고려하지 않은 단순한 정보제공에 그치고 있다.

추천시스템은 개인의 선호에 맞는 정보를 선택하여 제공한다는 점에서 전시회 내에서 관람객의

선호에 부합하는 부스를 추천하여 관람객의 니즈를 충족시키는 도구로 사용될 수 있다. 그러나 추천시스템 중 가장 성공적으로 평가 받는 협업 필터링은 실제 추천 과정에서 관람객의 방문 순서가 고려되지 않아 관람객의 선호 변화가 반영되지 않을 뿐만 아니라 관람객이 추천을 요청한 시점 뒤에 방문할 가능성이 낮은 부스를 추천하는 경우도 생기는 한계점이 존재한다. 따라서 본 연구에서는 순차 연관 규칙을 이용하여 관람객의 방문 순서를 고려하여 목표 관람객이 추천을 요청한 시점 이후에 방문할 가능성이 가장 높은 부스를 추천하는 방법론을 제안하였다.

제안한 방법론의 효과성을 검증하기 위해 실제 전시장에서의 데이터를 사용하여 기존 협업 필터링과의 성과를 비교하였다. 그 결과 전통적인 협업 필터링보다 본 연구에서 제안한 방법론이 정확도 면에서 더 높은 성과를 보였다. 특히 상위 1개의 부스를 추천하였을 경우의 정확도는 기존의 협업 필터링과 비교하여 큰 차이를 보였다. 따라서 작은 스크린(screen)을 가지고 있는 유비쿼터스 단말기에 적합하다. 또한, 추천에 사용되는 목표 관람객의 데이터가 3개일 때 성과가 가장 높게 나와 목표 관람객이 전시회를 방문한 초기에도 추천 서비스의 제공이 가능하다.

본 연구는 다음과 같은 한계점이 존재한다.

첫째, 순차 연관 규칙은 관람객들의 부스 방문 기록을 사용하여 연관 규칙을 생성하게 된다. 따라서 규칙 생성에 사용되는 데이터의 양이 증가할수록 정확한 규칙의 생성이 가능하다. 하지만, 전시 환경은 단기적이고 일시적으로 열리는 특성상 연관 규칙 생성에 사용될 수 있는 데이터의 한계가 있다. 특히, 전시회가 열린 초기에는 연관 규칙을 생성하는 데 사용할 데이터가 희박해 추천 시스템의 성과가 현저하게 저하될 수 있다.

둘째, 본 연구에서 사용된 순차 연관 규칙은 빈도 계산을 위해 고려되는 후보 선행 규칙의 순서는 고려하지 않는다. 또한 목표 관람객의 부스 방문 기록에서 목표 관람객의 선행 규칙을 생성하는 과정에서 규칙의 시점도 고려하지 않는다. 예를 들어, 후보 선행 규칙이 {B1, B2}이고 목표 관람객이 부스 방문 기록이 B2 → B3 → B1인 경우임에도 순차 연관 규칙은 목표 관람객의 부스 방문 기록에 후보 선행 규칙이 존재한다고 판단하여 신뢰도 정보를 제공한다. 따라서 목표 관람객의 현재 시점에 영향력이 약한 과거 규칙을 바탕으로 추천 정보를 제공하여 현재 시점의 추천의 성과를 저하시킬 가능성이 존재한다.

마지막으로, 성과 평가를 위한 실제 전시회의 부스 방문 기록 데이터는 상품의 수, 즉 부스의 수가 24개로 적은 편이다. 또한, 부스의 수가 적음에 따라 관람객들의 부스 방문 패턴이 몇 가지의 패턴으로 대표된다. 최근 전시회의 대형화로 인해 추천시 고려해야 하는 부스의 수가 증가하였으며 관람객들의 선택 폭이 넓어짐에 따라 관람객의 방문양상도 개별 관람객 별로 다양하게 존재한다. 따라서 추천 시스템의 성과를 평가하기 위한 데이터로는 더 많은 부스를 포함한 데이터가 필요하다.

따라서 향후 연구에서는 위와 같은 제한점을 해결하기 위해 먼저, 더 넓은 전시 공간을 갖춘 전시회를 대상으로 데이터를 수집하여 본 연구에서 제시한 방법론의 성과를 검증하고, 전시 환경에서 가지는 한계점을 해결하기 위해 방법론을 정교화 할 것이다.

참고문헌

김재경, 김혜경, 최일영, “유비쿼터스 환경에서 개

체간의 자율적 협업에 기반한 추천방법 개발”, *지능정보연구*, 15권 1호(2009), 31~50.

김재경, 안도현, 조운호, “인터넷 쇼핑몰을 위한 데이터 마이닝 기반 개인별 상품 추천방법론의 개발”, *지능정보연구*, 9권, 3호(2003), 177~191.

이연남, 권오병, “추천시스템 연구의 개발추세 동향”, *지능정보연구*, 14권, 2호(2008), 63~82

이희찬, 한진영, “전시관련 서비스 품질의 중요도-실행도 불일치가 전시관람 행동에 미치는 영향”, *한국관광학회*, 29권, 3호(2005), 165~184.

전정호, 이경진, “모바일 RFID에 기반한 유비쿼터스 전시공간 비즈니스 모델 설계 및 사례 연구”, *지능정보연구*, 14권, 4호(2009), 47~68.

Agrawal, R., T. Imielinski and A. Swami, “Mining association rules between sets of items in large databases”, *In Proc. 1993 ACM-SIGMOD Int. Conf. Management of Data*, (1993), 207~216.

Amer-Yahia, S., L. Lakshmanan and C. Yu, “Socialscope : Enabling information discovery on social content sites.”, *CIDR*, 2009.

Bellizi, J. and D. Lipps, “Managerial Guidelines for Trade Show Effectiveness”, *Industrial Marketing Management*, Vol.13, No.1(1984), 49~52.

Billsus, D. and M. J. Pazzani(1998), “Learning collaborative information filters”, *In Proceedings of the 15th International Conference on Machine Learning*, (2000), 46~54.

Blythe, J., “Visitor and exhibitor expectations and outcomes at trade exhibitions”, *Marketing Intelligence and Planning*, Vol. 17, No.2(1999), 100~108.

Harms, S. K., J. Deogun and T. Tadesse, “Discovering Sequential Association Rules with Constraints and Time Lags in Multiple Sequences”, *In Proceedings of the 2002 International Symposium on Methodologies for*

- Intelligent Systems*, (2002), 432~441.
- Herlocker, J., J. Konstan, L. Terveen and J. Riedl, "Evaluating collaborative filtering recommender systems", *ACM Transactions on Information Systems*, Vol.22, No.1(2004), 5~53.
- John, Jr. F., Tanner, B. C. Lawrence, and V. P. Thomas, "A Learning Model of Trade show Attendance", *Journal of Convention and Exhibition Management*, Vol.3, No.3(2001), 3~26.
- Kim, D. H., V. Atluri, M. Bieber, N. Adam, and Y. Yesha, "A clickstream based collaborative filtering personalization model : towards a better performance", *In Proceedings of the 6th annual ACM international workshop on Web information and data management*, (2004), 88~95.
- Lu, Z., D. Agarwal, and I. Dhillon, "A spatio-temporal approach to collaborative filtering", *In Proceedings of the 3rd ACM International Conference on Recommender Systems*, 2009.
- Mobasher, B., R. Cooley, and J. Srivastava, "Automatic Personalization Based on Web Usage Mining", *Communication of the ACM*, Vol. 43, No. 8, 142~151.
- Ng, R., L. S. Lakshmanan, J. Han and A. Pang, "Exploratory Mining and Pruning Optimizations of Constrained Associations Rules", *In Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*. 1998.
- Sandra, L. M., "The Art Of The Show", *LAEM Foundation*. 2002.
- Sarwar, B. M., G. Karypis, J. A. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e commerce", *In Proceedings of the 2nd ACM conference on Electronic commerce*, (2000), 158~167.
- Srikant, R., Q. Vu, and R. Agrawal, "Mining Association Rules with Item Constraints", *In Proceedings of the Third International Conference on Knowledge Discovery and Data Mining*, (1997), 67~73.
- Zaki, M., "Sequence mining in Categorical Domains : Incorporating Constraints", *In Proceedings of the Ninth International Conference on Information and Knowledge Management*, (2000), 422~429.

Abstract

Personalized Exhibition Booth Recommendation Methodology Using Sequential Association Rule

Hyun Sil Moon* · Min Kyu Jung* · Jae Kyeong Kim* · Hyea Kyeong Kim*

An exhibition is defined as market events for specific duration to present exhibitors' main product range to either business or private visitors, and it also plays a key role as effective marketing channels. Especially, as the effect of the opinions of the visitors after the exhibition impacts directly on sales or the image of companies, exhibition organizers must consider various needs of visitors. To meet needs of visitors, ubiquitous technologies have been applied in some exhibitions. However, despite of the development of the ubiquitous technologies, their services cannot always reflect visitors' preferences as they only generate information when visitors request. As a result, they have reached their limit to meet needs of visitors, which consequently might lead them to loss of marketing opportunity. Recommendation systems can be the right type to overcome these limitations. They can recommend the booths to coincide with visitors' preferences, so that they help visitors who are in difficulty for choices in exhibition environment. One of the most successful and widely used technologies for building recommender systems is called Collaborative Filtering. Traditional recommender systems, however, only use neighbors' evaluations or behaviors for a personalized prediction. Therefore, they can not reflect visitors' dynamic preference, and also lack of accuracy in exhibition environment. Although there is much useful information to infer visitors' preference in ubiquitous environment (e.g., visitors' current location, booth visit path, and so on), they use only limited information for recommendation.

In this study, we propose a booth recommendation methodology using Sequential Association Rule which considers the sequence of visiting. Recent studies of Sequential Association Rule use the constraints to improve the performance. However, since traditional Sequential Association Rule considers the whole rules to recommendation, they have a scalability problem when they are adapted to a large exhibition scale. To solve this problem, our methodology composes the confidence database before recommendation process. To compose the confidence database, we first search preceding rules

* School of Management and Management Research Institute, Kyunghee University

which have the frequency above threshold. Next, we compute the confidences of each preceding rules to each booth which is not contained in preceding rules. Therefore, the confidence database has two kinds of information which are preceding rules and their confidence to each booth. In recommendation process, we just generate preceding rules of the target visitors based on the records of the visits, and recommend booths according to the confidence database. Throughout these steps, we expect reduction of time spent on recommendation process.

To evaluate proposed methodology, we use real booth visit records which are collected by RFID technology in IT exhibition. Booth visit records also contain the visit sequence of each visitor. We compare the performance of proposed methodology with traditional Collaborative Filtering system. As a result, our proposed methodology generally shows higher performance than traditional Collaborative Filtering. We can also see some features of it in experimental results. First, it shows the highest performance at one booth recommendation. It detects preceding rules with some portions of visitors. Therefore, if there is a visitor who moved with very a different pattern compared to the whole visitors, it cannot give a correct recommendation for him/her even though we increase the number of recommendation. Trained by the whole visitors, it cannot correctly give recommendation to visitors who have a unique path. Second, the performance of general recommendation systems increase as time expands. However, our methodology shows higher performance with limited information like one or two time periods. Therefore, not only can it recommend even if there is not much information of the target visitors' booth visit records, but also it uses only small amount of information in recommendation process. We expect that it can give real-time recommendations in exhibition environment.

Overall, our methodology shows higher performance ability than traditional Collaborative Filtering systems, we expect it could be applied in booth recommendation system to satisfy visitors in exhibition environment.

Key Words : Sequential Association Rule, Exhibition Booth Recommendation, Recommendation System, Recommendation

저자 소개



문현실

경희대학교에서 경영학 학사를 취득하였으며, 동 대학원에서 경영정보전공 석사과정에 재학 중이다. 주요 관심분야로는 데이터마이닝, 추천시스템, 사회연결망분석, 복잡계 시스템 등이다.



정민규

경희대학교 전자공학 학사를 취득하였으며, 현재 동 대학 경영학과 일반대학원 경영정보 전공 석사 과정에 재학 중이다. 주요 관심분야는 추천시스템, 데이터마이닝, 유비쿼터스 서비스 등이다.



김재경

서울대학교에서 산업공학 학사, 한국과학기술원에서 경영정보시스템 전공으로 석사 및 박사학위를 취득하였으며 현재 경희대학교 경영대학 교수로 재직하고 있다. 미국 미네소타 주립대학교, 그리고 텍사스 주립대학교(달라스)에서 교환교수를 역임하였다. 주요 관심분야로는 비즈니스 인텔리전스, 추천시스템, 유비쿼터스 서비스 등이다. 다수의 국내외 학술지에 논문을 게재하였으며, 한국지능정보시스템학회, 저탄소녹색성장국민포럼 그린IT분과 위원, 경희대학교 경영대학 BK21 사업단장, Information Technology and Management(SSCI) AE(Associate Editor)를 역임중이다.



김혜경

현재 경희대학교 경영대학에서 연구교수로 재직하고 있다. 경희대학교 물리학과에서 학사, 일반대학원 경영학과에서 MIS 전공으로 석사학위와 박사학위를 취득하였다. 주요 관심분야는 고객관계관리, 상품추천시스템, 사회연결망분석, 복잡계 시스템 등이며, International Journal of Information Management, IEEE Transactions on Services Computing, Expert Systems, Expert Systems With Applications, LNCS, LNAI 등 에 관련논문을 게재하였다.