

유비쿼터스 환경에서 사용자의 직접 피드백을 최소화하기 위한 컨텍스트 마이닝 시스템 설계

The Design of A Context Mining System to Minimize Users' Direct-Feedback in Ubiquitous Environment

최영환, 이상용

공주대학교 컴퓨터공학과, 공주대학교 컴퓨터공학부

Young-Hwan Choi, Sang-Yong Lee

Dept. of Computer Engineering, Division of Computer Science & Engineering,

Kongju National University

E-mail : cyhmad@kongju.ac.kr

요 약

현재 유비쿼터스 환경에서 대부분의 시스템이 개인화된 추천 서비스를 위한 컨텍스트 인식 과정에서 사용자의 직접 피드백을 받는 경우가 많다. 다양한 서비스가 사용자 주변에 존재한다고 하더라도 사용자가 서비스를 받기 위해 직접 피드백을 하는 경우가 많아지면 invisible service를 받을 수 없게 된다. 본 논문에서는 마이닝 기법을 기반으로 사용자의 프로파일 생성과 갱신, 선호도를 예측하여 효율적인 서비스를 제공하는 컨텍스트 마이닝 시스템을 제안한다. 본 시스템에서는 초기프로파일을 생성할 때만 사용자의 직접 피드백을 이용하고, 사용자 프로파일의 갱신과 선호도 예측, 추천 등 컨텍스트 마이닝 과정에서는 사용자의 행동과 사용자 유사한 그룹의 선호도, 그리고 사용자의 주변 환경과 같은 컨텍스트 정보를 이용하여 직접 피드백을 최소화한다.

1. 서론

최근 새로운 IT 패러다임인 유비쿼터스 컴퓨팅의 발전으로 사용자들이 필요한 정보나 서비스를 언제 어디서나 즉시에 제공 받을 수 있게 되었다. 하지만, 현재 서비스 시스템의 대부분이 사용자에게 개인화된 추천 서비스를 하기 위해 컨텍스트를 얻을 때 사용자의 직접 피드백을 받는 경우가 많다. 예로 고객이 쇼핑몰에 들어갔을 때 PDA에 구매하고자 하는 물품을 직접 입력하고, 시스템은 상품정보를 고객에게 제공 해주거나, 영화관에 갔을 때 원하는 장르의 영화를 고객이 선택하고 시스템은 서비스를 제공하는 방법을 사용하고 있다.

다양한 서비스가 사용자 주변에 존재한다고 하더라도 사용자가 서비스를 받기 위해 직접 원하는 항목을 입력하고 서비스에 대해 평가를 하는

직접 피드백이 많아지면 유비쿼터스 컴퓨팅의 목적인 invisible service를 할 수 없게 된다. 이러한 문제를 해결하기 위해 사용자가 인식하지 못하는 상태에서 사용자의 정확한 기호와 성향뿐만 아니라 주변 환경까지 고려한 컨텍스트 인식 시스템이 필요하다.

본 논문에서는 데이터마이닝 기법을 기반으로 효율적인 컨텍스트 인식을 위한 컨텍스트 마이닝 시스템을 제안한다. 초기 프로파일을 생성할 때만 사용자의 직접 피드백을 이용하고, 사용자 프로파일의 갱신과 선호도 예측, 추천 등 컨텍스트 마이닝 과정에서는 사용자의 행동과 사용자 유사한 그룹의 선호도, 그리고 사용자의 주변 환경과 같은 간접 피드백을 이용한다.

2. 관련연구

2.1 컨텍스트

유비쿼터스 컴퓨팅에서 사용되는 컨텍스트의 개념은 아직까지 통일된 정의가 없는 상태이며, 많은 연구자들이 독자적인 정의를 바탕으로 객관화된 컨텍스트의 개념을 정의하기 위하여 꾸준히 시도하고 있다. Schilit[1] 등은 사용자의 위치, 사용자 주변에 있는 사람 정보, 그리고 사용 가능한 자원 등의 정보를 컨텍스트라고 정의하였다. 이것은 사용자의 환경이 변화하여도 일관성 있는 컨텍스트의 개념이 적용될 수 있는 특징을 나타낸다. Dey[2] 등은 기존 정의들을 종합하여 컨텍스트를 사용자와 응용서비스 사이의 상호작용을 위해 필요한 사용자, 장소, 대상물 등의 개체 상태를 나타내는 정보라고 정의하였으며, 이 개념이 최근 여러 사람들에 의해 많이 참조가 되고 있다. Jang[3] 등은 컨텍스트란 5W1H : Who, What, Where, When, Why, How 이며, 응용 서비스에 따라 5W1H의 조합으로 나타난다고 정의하였다.

2.2 컨텍스트 인식 시스템

컨텍스트가 사용자의 상황을 나타내는 정보라고 하면, 컨텍스트 인식 시스템은 이러한 정보를 기반으로 적절한 서비스를 제공하는 시스템을 의미한다. 컨텍스트 인식 시스템들은 크게 두 종류로 나뉜다. 첫째, 주변 환경의 변화를 감지하여 컨텍스트를 파악하고, 컨텍스트 정보에 따라 적절한 서비스를 제공하는 것이다.[1][2][4] 둘째, 컨텍스트에 맞추어 시스템의 실행 조건이나 주변 환경을 스스로 변경하는 것이다.

현재 유비쿼터스 컴퓨팅에서 컨텍스트 인식 시스템과 응용 서비스에 대한 개발이 활발하게 진행되고 있다. 사무 환경, 가정 환경, 공동 작업 환경, 교육 환경, 실내 환경 등의 분야에서 연구가 많이 진행되고 있으며, 사용자의 컨텍스트를 직접 피드백이 아닌 간접 피드백으로부터 얻어낸다. 하지만, 오프라인 마켓이나 극장, 식당 등 사용자에게 개인화된 서비스를 추천하는 분야에서는 사용자의 직접 피드백으로부터 컨텍스트를 얻어내는 경우가 대부분이다.

2.3 데이터 마이닝

데이터 마이닝이란 방대한 양의 데이터에 함축적으로 들어 있는 지식이나 패턴을 찾아내는 기술이다. 데이터 마이닝을 분류하면 친화도나 패턴을 찾아내는 연관규칙[5], 시간이라는 개념이 포함되어 동시에 발생될 가능성이 큰 데이터 집합을 찾아내는 순차패턴, 그룹화 되어있는 객체 집합에 대한 분류규칙, 유사성이 높은 객체들을

그룹화하는 클러스터링, 유사 패턴을 찾아내는 유사성 탐색 등이 있다.

3. 컨텍스트 마이닝 시스템 설계

본 논문에서 제안하는 시스템은 사용자의 컨텍스트를 데이터 마이닝 기법으로 분석하여 사용자의 구매성향이나 구매패턴을 파악하여 적절한 상품을 실시간으로 추천한다. 먼저, 클러스터링 기법 중 협력적 필터링[6]을 사용하여 건물 내 각 매장의 상품과 사용자의 데이터를 분류하고 유사성이 높은 사용자와 상품을 그룹화하고, 연관규칙을 이용하여 추천 리스트를 제공한다.

이 시스템에서 사용하는 컨텍스트 항목은 초기 프로파일 생성시 이름, 성별, 생년월일을 사용하였고, 나머지는 ID, 구매이력, 체류시간을 사용하였다. 컨텍스트 마이닝 시스템의 구조는 그림 1과 같이 컨텍스트 인식 모듈, 추천 모듈, 상품 정보 DB, 사용자 프로파일 DB, 프로파일 매니저 에이전트(PMA)로 구성된다.

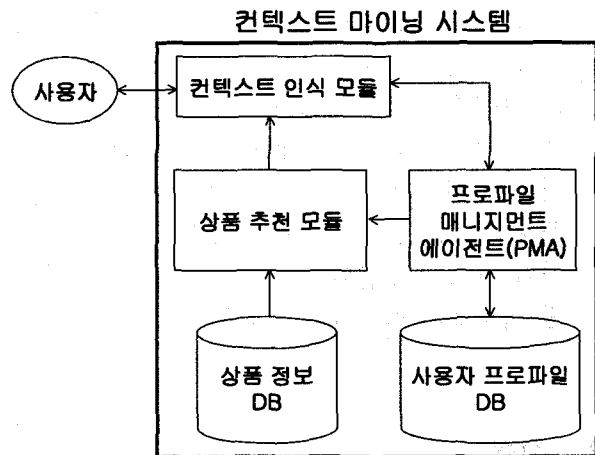


그림 1 컨텍스트 마이닝 시스템의 구조

컨텍스트 인식 모듈에서는 사용자가 카트를 가지고 매장에 들어오게 되면 카트에 부착되어 있는 PDA로부터 사용자의 아이디를 인식하여 PMA로 보내고, 체류시간을 얻어내기 위해 사용자의 위치를 계속 모니터링한다. PMA는 사용자프로파일을 생성하고 갱신하며, 사용자프로파일 DB로부터 해당 ID를 가지는 사용자의 추천목록을 사용자 프로파일 DB로부터 추천 모듈로 보낸다. 상품 추천 모듈은 추천 목록으로 상품 정보 DB에서 상품 정보를 컨텍스트 인식 모듈로 보내고, 컨텍스트 인식 모듈은 상품 정보를 PDA를 통해 사용자에게 추천해준다.

3.1 사용자 프로파일

초기 사용자 프로파일은 쇼핑몰에 처음 사용자가 등록을 하고 처음 상품을 구매하게 되면 PMA는 항목 중 이름, 성별, 생년월일과 구매이력을 기준으로 사용자 프로파일 DB로부터 같은 그룹에 속하는 사용자들의 프로파일을 사용하여 생성한다. 표 1은 사용자 프로파일 생성과 갱신에 사용된 컨텍스트 항목들이다.

ID	731026MHKD001
이름	홍길동
성별	Male
생년월일	1973-10-26
선호도	각 상품에 대한 선호도
구매이력	코드, 상품명, 구매일, 상품가격
체류시간	각 상품에 대한 체류시간

표 1 사용된 컨텍스트 항목

3.2 사용자의 선호도 계산

초기 프로파일을 기반으로 사용자의 상품에 대한 선호도를 계산하기 위해 사용자의 구매이력과 컨텍스트 인식모듈로부터 얻은 상품에 대한 체류시간을 사용한다. 사용자의 구매이력에는 각 상품에 대해 상품코드, 상품명, 구매일자, 상품가격 등의 정보가 저장되어 있다. 사용자의 체류시간은 해당하는 코너에 들어간 시간부터 나온 시간을 측정한다. 상품에 대한 체류시간이 일정 시간을 넘어가면 해당 상품에 관심이 있는 것으로 판단하여 가중치 값을 증가시켜준다. 또, 추천 모듈로부터 추천받은 상품을 사용자가 구매했을 경우에는 긍정적 피드백으로 인식하여 가중치를 증가시키고, 구매하지 않은 경우에는 부정적 피드백으로 인식하여 가중치를 감소시킨다.

그 이후 협력적 필터링을 이용하여 사용자들을 보다 유사한 취향을 가진 그룹으로 다시 그룹화하고, 그룹 내 사용자들의 컨텍스트를 이용하여 협력적 추천이 이루어지게 된다.

사용자 *a*와 사용자 *k*의 유사도 $w_{a,k}$ 의 계산은 식(1)의 Pearson 상관관계 계수식을 통해 계산한다. 유사도 값은 -1.0에서 1.0 사이의 값을 가지며, 큰 값일수록 유사도가 높음을 의미한다. 계산된 유사도를 식(2)에 적용하여 새로운 상품 *i*에 대한 선호도를 예측한다.

$$w_{a,k} = \frac{\sum_j (r_{a,j} - \bar{r}_a)(r_{k,j} - \bar{r}_k)}{\sqrt{\sum_j (r_{a,j} - \bar{r}_a)^2 \times \sum_j (r_{k,j} - \bar{r}_k)^2}} \quad \text{식(1)}$$

$$P_{a,i} = \bar{r}_a + \frac{\sum_k \{(r_{k,i} - \bar{r}_k) \times w_{a,k}\}}{\sum_k |w_{a,k}|} \quad \text{식(2)}$$

여기서 *j*는 사용자 *a*와 *k*가 동시에 선호도를 준 아이템을 의미하고 $r_{a,j}$ 는 사용자 *a*의 상품 *j*에 대한 선호도를 말하고, \bar{r}_a 는 사용자 *a*의 평균 선호도를 나타낸다.

3.3 사용자 프로파일 갱신

사용자의 선호도는 사용자의 구매이력과 체류시간, 같은 그룹에 속한 사람들의 선호도를 이용해 하루에 한 번 갱신하여 사용자 프로파일 DB에 저장한다.

사용자 프로파일의 갱신은 먼저 오랫동안 머문 상품에서의 선호도를 적용시키기 위해 각각의 선호도를 0.1씩 증가시킨다. 다음으로 사용자의 구매이력으로부터 추천 상품에 대한 구매는 긍정적 피드백으로, 추천 상품에 대한 비구매는 부정적 피드백을 구분하여 각각의 상품에 대한 가중치를 곱한 값을 계산한 후 각각의 max 값을 이용하여 정규화하여 갱신할 값을 계산한다. 마지막으로 프로파일에 저장되어 있지 않은 내용은 새로이 추가하고 이미 존재하는 내용에 대해서는 기존의 선호도에 갱신 값을 더한다. 갱신 후의 선호도가 0일 때에는 프로파일에서 삭제한다.

$$V_{a,j} = \sum_{j \in F(+)} pr_{a,j} + \sum_{j \in F(-)} nr_{a,j}$$

$$V_{a,j} = \frac{V_{a,j}}{\max(V_{a,j}, j)} \quad \text{식(3)}$$

$$r_{a,j} = r_{a,j} + V_{a,j} \cdot j$$

a : 사용자

F(+) : 사용자가 긍정적 피드백을 준 상품의 집합

F(-) : 사용자가 부정적 피드백을 준 상품의 집합

j : 상품

$pr_{a,j}$: 긍정적 피드백을 얻은 상품의 가중치

$nr_{a,j}$: 부정적 피드백을 얻은 상품의 가중치

갱신된 사용자 프로파일을 토대로 사용자가 관심을 가질만한 품목을 예측하고, 컨텍스트 인식 모듈로부터 사용자의 현재 위치를 파악하여 사용자의 주변에 있는 상품을 추천한다.

4. 결론 및 향후과제

본 논문에서는 유비쿼터스 환경에서 사용자의 직접피드백을 최소화하는 컨텍스트 마이닝 시스템

템을 설계하였다. 이 시스템에서는 컨텍스트를 분석하여 사용자의 취향에 적합한 상품을 추천하고, 사용자의 간접 피드백을 반영하여 동적으로 사용자 프로파일을 갱신하고 이를 추천결과에 반영한다. 사용된 컨텍스트 항목은 이용 가능한 모든 개체들의 정보를 뜻하며 제안된 시스템에서는 사용자의 위치와 구매이력, 체류시간 등의 간접 피드백을 사용함으로써 더 이상의 직접 피드백은 요구되지 않는다.

향후 본 시스템을 구현하고 추천의 정확도를 보다 향상시키기 위해 사용자의 이동경로를 예측하여 사용자 프로파일 갱신에 반영한다면 보다 나은 추천 결과를 얻을 수 있을 것이다.

6. 참고문헌

- [1] B. Schilit, N. Adams, and R. Want, "Context-Aware Computing Applications," 1st International Workshop on Mobile Computing Systems and Application., pp.85-90, 1994.
- [2] A. K. Dey, and G. D. Abowd, "Towards a Better Understanding of Context and Context-Awareness," In Proceedings of the CHI 2000 Workshop on The What, Who, Where, When, and How of Context-Awareness (The Hague, Netherlands), Apr. 2000.
- [3] 장세이, 우운택, "유비쿼터스 컴퓨팅 환경을 위한 컨텍스트 기반 애플리케이션 구조", 한국정보과학회 HCI 논문집, 제2권, pp.346-351, 2003.
- [4] J. Pascoe, "Adding Generic Contextual Capabilities to Wearable Computers," 2nd International Symposium on Wearable Computers, pp. 92-99, 1998.
- [5] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Analysis of Recommendation Algorithms for E-Commerce," *The ACM E-Commerce 2000 Conference*, 2000.
- [6] Lyle H. Ungar and Dean P. Foster. "Clustering Methods for Collaborative Filtering," *Proceeding of the 1998 Workshop on Recommendation Systmes*, pp.114-129, 1998.