

# Collaborative Filtering for Credit Card Recommendation based on Multiple User Profiles

Won Cheol Lee\* · Hyoup Sang Yoon\*\* · Seok Bong Jeong\*\*\*†

\*School of Business, Kyungil University,

\*\*Business School, Daegu Catholic University

\*\*\*Dep. of Railway, Kyungil University

## 신용카드 추천을 위한 다중 프로파일 기반 협업필터링

이원철\* · 윤협상\*\* · 정석봉\*\*\*†

\*경일대학교 경영학과

\*\*대구가톨릭대학교 경영학부

\*\*\*경일대학교 철도학과

Collaborative filtering, one of the most widely used techniques to build recommender systems, is based on the idea that users with similar preferences can help one another find useful items. Credit card user behavior analytics show that most customers hold three or less credit cards without duplicates. This behavior is one of the most influential factors to data sparsity. The 'cold-start' problem caused by data sparsity prevents recommender system from providing recommendation properly in the personalized credit card recommendation scenario. We propose a personalized credit card recommender system to address the cold-start problem, using multiple user profiles. The proposed system consists of a training process and an application process using five user profiles. In the training process, the five user profiles are transformed to five user networks based on the cosine similarity, and an integrated user network is derived by weighted sum of each user network. The application process selects k-nearest neighbors (users) from the integrated user network derived in the training process, and recommends three of the most frequently used credit card by the k-nearest neighbors. In order to demonstrate the performance of the proposed system, we conducted experiments with real credit card user data and calculated the F1 Values. The F1 value of the proposed system was compared with that of the existing recommendation techniques. The results show that the proposed system provides better recommendation than the existing techniques. This paper not only contributes to solving the cold start problem that may occur in the personalized credit card recommendation scenario, but also is expected for financial companies to improve customer satisfactions and increase corporate profits by providing recommendation properly.

**Keywords** : Recommender System, Collaborative Filtering, Multiple Profiles, Credit Card, Cold Start

## 1. 서론

신용카드는 소비자에게 편리한 휴대와 신용공여를 통해 미래의 소득으로 현재의 지출을 가능케 하는 현금대체 지급결제수단으로써 이용도가 높은 금융상품이다. 2016년 말 현재, 신용카드 발급매수는 20,412만 매(체크카드 포함)이며, 이용실적은 746조 원으로 소득공제율이 확대 시행된 2010년 이후 지속적으로 증가되고 있다[9].

신용카드 이용도 증가에도 불구하고 시장상황은 카드사에 우호적이지 않다. 2012년 여신금융업법 개정에 따른 가맹점 수수료율 인하로 인해 수익성이 악화되고 있고, 카드사 간 경쟁심화로 제휴사와의 협상력이 약화되어 마케팅 비용이 증가하고 있다. 최근에는 스마트기기 도입으로 간편결제 서비스와 같은 신용카드를 대체할 수 있는 지급결제 수단의 등장으로 경쟁이 심화되고 있어 신용카드 이용액의 증가 추세는 둔화되고 있다[7]. 또한, 타 업종의 지급결제 시장 진입에 대한 장벽도 무너지고 있어 한정된 고객을 유치하기 위한 경쟁이 치열해지고 있다.

카드사는 더 좋은 혜택과 다양한 서비스를 원하고 있는 고객의 요구사항을 충족하기 위한 마케팅 활동에 많은 노력을 기울이고 있다. 일반적인 카드사의 마케팅 활동은 각종 포인트 제공, VIP서비스, 편의서비스, 할인혜택, 결제서비스, 카드 자동 납부 등 각종 혜택과 편의성을 갖춘 카드 상품 개발과 각종 이벤트 실시, 광고, 프로모션 등의 다양한 방법으로 구성되어 있다[12]. 신한카드의 7개의 전업카드사는 카드모집인의 오프라인 영업, TM(Tele marketing), DM(Direct marketing)방식을 주로 활용하였으나, 최근에는 인터넷의 발달과 사용인구의 증가로 카드사 홈페이지를 통해 맞춤형 카드 찾기, 내게 맞는 카드 찾기 등의 온라인 마케팅을 강화하고 있다. 농협 외 10개 경영은행계 카드사도 영업점 내점고객의 선호도 및 속성을 파악하여 적합한 카드를 추천하는 마케팅 형태를 취하고 있다.

카드사의 공통적인 전략은 고객의 다양한 요구를 만족시킬 수 있는 고객 맞춤형 카드를 출시하여 가입 고객 수와 카드 이용액을 증대시키는 것이다. 이러한 전략에 의해 카드사 별로 많은 수의 카드 상품을 보유하고 있다. 그러나 상담직원이 다양한 카드의 상세한 특성과 불특정 다수의 고객성향을 짧은 시간에 파악한다는 것은 쉽지 않기 때문에 고객 맞춤형 카드 추천에 어려움을 겪고 있다.

한편, 인터넷과 IT기술의 발전으로 도서 및 상품 추천 등과 같은 온라인 쇼핑 분야에서는 가용한 고객데이터가 다양해짐에 따라 추천 시스템의 구축과 관련된 연구가 활발하게 이루어지고 있다[6, 11, 13, 14, 15, 16, 17, 19, 21, 22]. 추천 시스템 구축을 위해 가장 널리 사용되고

있는 협업필터링(collaborative filtering)은 비슷한 취향을 갖는 타사용자의 평가를 활용하여 추천 대상을 선정하는 기법이다. 대표적으로 아마존, 구글, 넷플릭스, 티모, 그리고 야후와 같은 선도 기업들이 이러한 기술을 통해 개인화된 추천시스템을 운영하고 있다[2, 3, 4, 8, 18, 23].

본 연구에서는 협업필터링을 활용한 고객 맞춤형 신용카드 추천시스템 개발을 목적으로 하고 있다. 신용카드 추천은 일반적인 온라인 쇼핑과는 다르게 개인별 신용카드 보유 개수가 대략 3개 이하로, 반복적인 구매나 타 상품(카드)에 대한 수요가 지속적으로 발생하지 않는다는 특징을 가지고 있다. 따라서 협업필터링을 이미 활용하고 있는 선도 기업의 성공적인 적용 사례와는 달리 신용 카드 추천시스템을 개발하기 위해서는 협업필터링의 약점으로 지적되고 있는 콜드스타트(cold start) 문제[25]를 해결할 수 있는 기법이 필요하다.

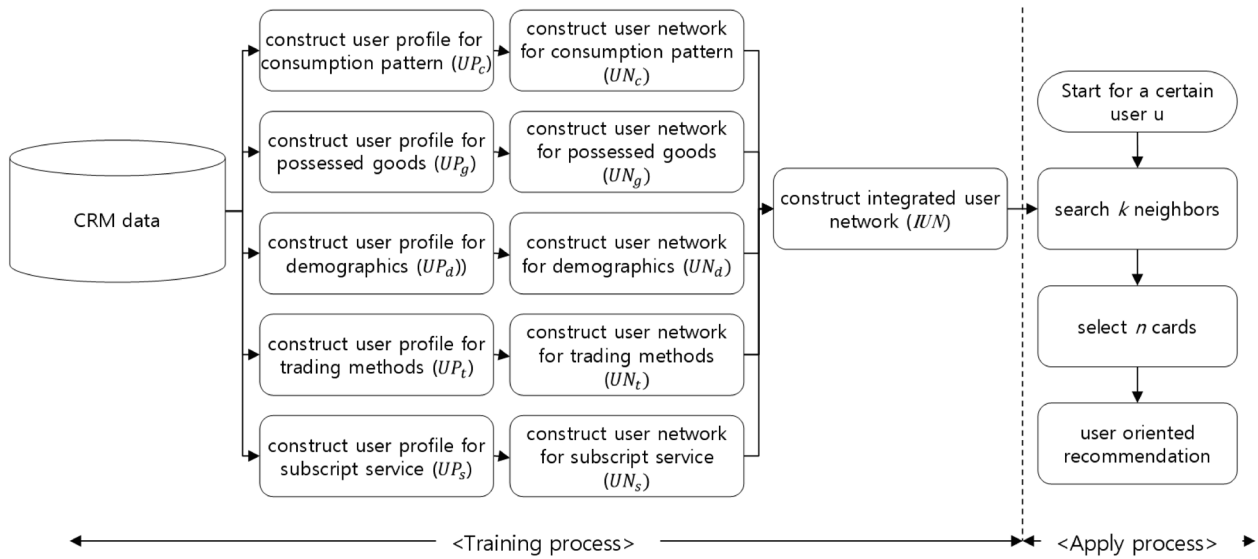
콜드스타트는 데이터가 충분히 축적되기 전까지 추천시스템의 성능이 저조하게 나타나는 현상을 의미하며 데이터 희박성(data sparsity)에 의해 발생한다[25]. 이러한 데이터 희박성은 다양한 종류의 아이템을 다수 취급하는 경우 전체 아이템 중에서 사용자가 평가하는 아이템의 종류가 극히 제한적인 경우 발생하게 된다. 또한, 새로운 사용자에게 아이템을 추천하는 경우나 새로운 아이템이 추가된 경우에도 데이터 희박성 문제가 발생할 수 있다.

본 연구에서는 신용카드 추천시스템의 콜드스타트 문제를 해결할 수 있는 다중 고객프로파일 기반 협업필터링 추천기법을 제안하고자 한다. 일반적인 협업 필터링 기법에서 각 아이템에 대한 사용자의 선호도나 구매내역 등의 단일 프로파일을 사용하는 것과는 달리, 본 연구에서는 고객의 금융거래 정보를 통해 수집할 수 있는 다양한 정보를 기반으로 다중 고객프로파일을 구성한다. 이러한 다중 고객프로파일은 고객의 신용카드에 대한 선호도나 사용내역 등의 정보 없이도 유사고객을 파악하는 것을 가능하게 하고 궁극적으로 콜드스타트 문제를 해결하는데 도움을 준다.

본 논문의 구성은 다음과 같다. 제 2장에서는 제안하고자 하는 추천기법에 대해 소개한다. 제 3장에서는 국내 A은행의 데이터를 활용하여 제안된 기법의 성능을 평가하고, 마지막으로 제 4장에서는 본 연구의 의의 및 추후 연구에 대해 서술한다.

## 2. 고객맞춤형 신용카드 추천 기법

본 연구에서 제안하는 고객맞춤형 신용카드 추천 기법은 크게 (1) 다수의 고객프로파일을 기반으로 통합 고객 네트워크(integrated user network, IUN)를 구성하는 부분



<Figure 1> Framework for the Proposed Scheme

과, (2) *IUN*을 활용한 고객 별 맞춤형 카드 도출부분으로 구성된다(<Figure 1> 참조).

제안된 기법은 일반적인 머신러닝 방법론에 따라 학습과정과 적용과정으로 이루어지는데, 학습과정에서는 고객 정보의 변경에 따라 주기적으로 각 고객프로파일을 갱신하고 고객 네트워크를 구성한다. 이어 적용과정에서는 은행 방문 고객을 대상으로 해당 고객과 가장 유사한 *k*명의 이웃고객을 선별하고, 이웃고객들이 가장 많이 사용하는 *n*개의 카드를 도출하여 카드 추천에 활용한다.

각 과정에 대한 상세한 설명은 다음절부터 제시되어 있다.

2.1 학습과정 : 통합 고객 네트워크 *IUN* 구성

학습과정은 *IUN*을 도출하는 과정으로, 이는 (1) 고객의 거래정보로부터 개별 고객프로파일을 구성하는 부분과, (2) 구성된 개별 고객프로파일로부터 개별 고객 네트워크를 구성하는 부분, 그리고 (3) 개별 고객 네트워크를 통합

하여 *IUN*을 구성하는 부분으로, 총 3단계로 이루어진다.

학습과정의 첫 번째 단계인 고객프로파일 구성은 금융 기업의 고객데이터베이스로부터 획득할 수 있는 고객 정보를 바탕으로 이루어진다. 본 연구를 위하여 국내 A은행의 신용카드 고객 데이터를 활용하였다. A은행은 신용카드 연간 매출액 5조 8천억 원, 총 회원 수 180만 명, 150여 종의 카드 상품을 취급하는 국내 중규모의 금융회사이다. 본 연구에서 활용한 데이터를 기반으로 5종의 고객프로파일을 구성하였으며, 그 내역은 <Table 1>과 같다.

- 소비패턴 프로파일(user profile for consumption pattern, *UP<sub>c</sub>*) : 일반한식, 주유소, 스포츠, 병원, 홈쇼핑 외 75개 대표업종에 대한 고객의 1년 간 소비 금액을 표시하여, 고객의 주요 소비패턴을 알 수 있다. 소비패턴 프로파일은, <Figure 2>와 같이 2-mode 매트릭스로 표현된다. 예를 들어, <Figure 2>의 고객 *u1*은 1년 동안 한식에서 30만 원, 스포츠에 40만 원, 홈쇼핑에 50만 원을 지출하는 소비패턴을 보여주고 있다.

<Table 1> Attributes of Profiles

name of profile	Construct user profile for consumption pattern( <i>UP<sub>c</sub></i> )	Construct user profile for possessed goods( <i>UP<sub>g</sub></i> )	Construct user profile for demographics( <i>UP<sub>d</sub></i> )	Construct user profile for trading methods( <i>UP<sub>t</sub></i> )	Construct user profile for subscript service( <i>UP<sub>s</sub></i> )
no. of attributes	80	10	2	6	6
attributes	korea food, filling station, sports, hospital, home shopping, except 75 unit type of business	deferred savings, installment savings, any funds, installment funds, ELD, pension, loan, banca, ISA, trust	gender, age	automatic transfer, teller, internet, mobile, phone, ATM	internet banking, smart banking, phone banking, automatic transfer the apartment management fee, drawing transfer, SMS service

	korea food	filling station	sports	hospital	home shopping
u1	30	0	40	0	50
u2	10	20	0	0	0
u3	0	30	20	0	0
u4	0	15	0	35	0
u5	0	0	15	20	10

<Figure 2> User Profile for Consumption Pattern

- 보유상품 프로파일(user profile for possessed goods,  $UP_g$ ) : 보유상품 프로파일은 해당 금융기업의 판매 상품 중, 각 고객이 보유하고 있는 금융상품의 현황을 나타내고 있다. 이 프로파일의 각 열은 거치식, 적립식, 수익증권 등 금융기업이 취급하고 있는 10개의 금융상품을 표현하고 있다. 각 셀( $i, j$ )의 값은 0과 1의 이진수로 구성되는데, 고객  $i$ 가 상품  $j$ 를 보유하고 있는 경우 1, 아니면 0의 값을 갖는다. 이러한 보유상품 프로파일을 통해 고객의 금융상품 가입성향을 알 수 있다. <Figure 3>의 예에서 고객  $u1$ 은 거치식, 수익증권, 방카상품은 보유하고, 적립식과 대출상품은 보유하고 있지 않음을 보여준다.

	deferred savings	installment savings	funds	loan	banca
u1	1	0	1	0	1
u2	1	1	0	0	0
u3	0	0	1	1	0
u4	0	0	1	1	0
u5	1	0	0	1	0

<Figure 3> User Profile for Possessed Goods

- 인구통계학 프로파일(user profile for demographics,  $UP_d$ ) : A은행에서 획득할 수 있는 인구통계학 정보는 성별, 나이 등의 2개 속성으로 이루어지며, 성별은 남자와 여자로 구분되고, 나이는 10대부터 90대까지 분포되어 있다.
- 거래수단 프로파일(user profile for trading methods,  $UP_t$ ) : 거래수단 프로파일은 각 고객이 금융 업무를 수행하기 위하여 이용하는 수단을 표현하는 프로파일이다. 일반적인 금융업무 거래수단으로는 자동이체, 창구, 인터넷, 모바일, 폰, ATM 등이 있다. 거래수단 프로파일의 각 셀은 해당 고객이 각 이용수단의 1년간 이용건수를 나타내고 있으며, 이를 통해 고객의 금융업무 주 이용수단과 빈도를 알 수 있다.
- 가입서비스 프로파일(user profile for subscript service,  $UP_s$ ) : 가입서비스 프로파일은 인터넷뱅킹, 스마트뱅킹, 폰뱅킹, SMS 문자알림 서비스 등의 가입여부를 나타내는 이진행렬로 구성되어 사용자의 비대면 서비스 업무 이용선호도를 알 수 있다.

다음으로 학습과정의 두 번째 단계는 앞서 구성된 각 고객프로파일을 기반으로 1-mode의 고객 네트워크(user network)로 변환하는 과정이다. 본 연구에서는 아래와 같이, 5개의 고객프로파일 각각에 대응되는 5개의 고객 네트워크를 구성한다.

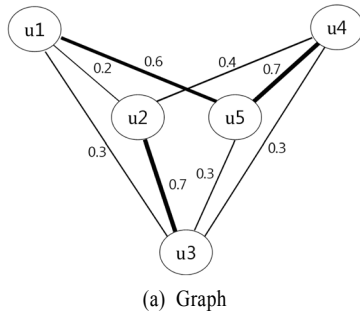
- 소비패턴 고객 네트워크 (user network for consumption pattern,  $UN_c$ )
- 보유상품 고객 네트워크 (user network for possessed goods,  $UN_g$ )
- 인구통계학 고객 네트워크 (user network for demographics,  $UN_d$ )
- 거래수단 고객 네트워크 (user network for trading methods,  $UN_t$ )
- 가입서비스 고객 네트워크 (user network for subscript service,  $UN_s$ )

위의 각 고객 네트워크는 고객프로파일에서 각 사용자 간의 유사도 계산을 통하여 구성할 수 있다. 기존 문헌에서 많이 사용하는 고객 유사도로는 피어슨상관계수 [19], 자카드계수[20], 코사인계수[10, 24] 등이 있다. 본 연구에서는 협업필터링에서 많이 사용되는 코사인 계수를 사용한다. 코사인 계수는 내적공간의 두 벡터 간 각도의 코사인 값을 이용하여 벡터간의 유사 정도를 측정하는 방법으로 식 (1)과 같이 표시되며, 여기서  $u_l$ 와  $u_m$ 는 사용자  $l$ ,  $m$  각각의 프로파일 벡터를 의미한다. 유사도 결과 값은 0부터 1사이의 값으로 나타나는데, 1에 가까울수록 비슷한 성향을 가지고 0이면 다른 성향을 지니고 있다는 것을 의미한다[10, 24]. 사용자  $l$ 와  $m$ 간의 유사도를 나타내는 코사인계수  $w_{lm}$ 는 식 (1)과 같이 구할 수 있다.

$$\omega_{lm} = \cos(\vec{u}_l, \vec{u}_m) = \frac{\vec{u}_l \cdot \vec{u}_m}{\|\vec{u}_l\| \times \|\vec{u}_m\|} \quad (1)$$

다만, 동일한 사용자 간의 유사도(유사도 행렬의 대각 성분)는 1이지만 분석결과의 의미 해석을 위해 동일한 사용자 간의 유사도는 0으로 변환한다. 예를 들어, <Figure 2>의 소비패턴 프로파일을 기반으로 코사인 계수를 사용한 소비패턴 고객 네트워크를 구성하면 <Figure 4>과 같다.

<Figure 4>에 따르면, 고객  $u2$ 의 경우 고객  $u3$ 와의 유사도 값이 0.74로 가장 유사한 것을 알 수 있다. 이는 고객  $u2$ 와 소비패턴이 가장 유사한 고객이  $u3$ 임을 의미한다. 이와 마찬가지로 <Figure 3>의 보유상품 프로파일을 기반으로 보유상품 고객 네트워크를 구성한 예가 <Figure 5>에 나타나있다.

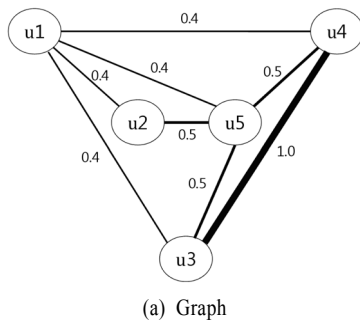


(a) Graph

	u1	u2	u3	u4	u5
u1	0.00	0.19	0.31	0.00	0.58
u2	0.19	0.00	0.74	0.35	0.00
u3	0.31	0.74	0.00	0.33	0.31
u4	0.00	0.35	0.33	0.00	0.68
u5	0.58	0.00	0.31	0.68	0.00

(b) Matrix

<Figure 4> User network for consumption pattern



(a) Graph

	u1	u2	u3	u4	u5
u1	0.00	0.41	0.41	0.41	0.41
u2	0.41	0.00	0.00	0.00	0.50
u3	0.41	0.00	0.00	1.00	0.50
u4	0.41	0.00	1.00	0.00	0.50
u5	0.41	0.50	0.50	0.50	0.00

(b) Matrix

<Figure 5> User Network for Possessed Goods

학습과정의 세 번째 단계는, 앞서 구성한 고객 네트워크를 기반으로 통합 고객 네트워크  $IUN$ 을 구성하는 과정이다. 본 연구에서는  $IUN$ 을 구축하기 위한 방법으로 Kim and Cho(2015)의 연구와 유사하게 단순평균 방식과 F1 기반의 가중평균 방식을 사용한다[15]. 단순평균 방식은 5개의 고객 네트워크에서 각 고객 간 유사도를 산술 평균하여  $IUN$ 의 유사도로 사용하는 방식이며, F1 기반의 가중평균방식은 각 고객 네트워크의 유사도를 가중 평균할 때 가중치로 각각의 네트워크의 추천 성능을 나타내는 F1 값을 사용한다.

한편, 기존의 많은 연구들이 추천 시스템의 성능을 평가하는 방법으로 정확도(Precision), 재현율(Recall) 및 F1

값을 사용해 왔다. 정확도는 추천 시스템이 추천한 아이템 중 실제 고객이 구매하는 아이템의 비율이며, 재현율은 고객이 구매한 아이템 중에서 추천 시스템의 추천이 적중시킨 비율이다. 그러나 이러한 정확도와 재현율은 시스템이 추천하는 아이템의 개수에 영향을 받는다. 따라서 최근의 추천시스템 연구에서는 식 (2)과 같이 정확도와 재현율을 조화 평균한 F1 값을 추천정확도의 평가 방법으로 사용한다[5, 6, 21, 26].

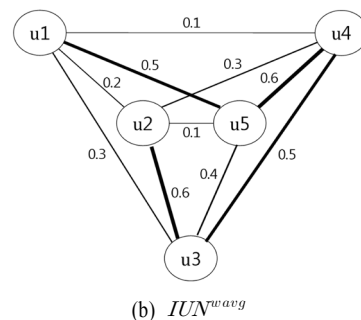
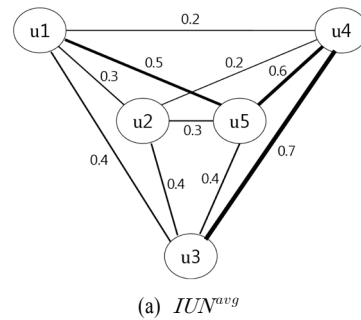
$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (2)$$

$i$  번째 고객 네트워크만을 기반으로 추천 상품을 도출했을 때의 F1 값을  $F1_i$ 라고 했을 때, 가중평균 방식으로 통합한  $IUN$ 은 다음 식 (3)과 같이 구할 수 있다.

$$IUN = \sum_i (F1_i \times UN_i) \quad \text{for } i \in \{c, g, d, t, s\} \quad (3)$$

단,  $c$  : 소비패턴,  $g$  : 보유상품,  $d$  : 인구통계학,  $t$  : 거래수단,  $s$  : 가입서비스

예를 들어, <Figure 4>와 <Figure 5> 두 개의 고객프로파일만 존재하고, 각각의 F1 값을 0.16과 0.05이라고 했을 때, 단순평균 및 가중평균으로 통합한 네트워크  $IUN^{avg}$  및  $IUN^{wavg}$ 는 <Figure 6>과 같다. <Figure 6>의 예에서  $IUN^{avg}$ 에서는  $u3$ 와 가장 유사한 고객은  $u4$ 지만,  $IUN^{wavg}$ 에서는  $u2$ 로,  $u3$ 의 이웃고객이 통합 네트워크 구축 방법에 따라 달라지는 것을 확인할 수 있다.



<Figure 6> Integrated User Network( $IUN$ )

## 2.2 적용과정 : IUN을 이용한 추천 카드 도출

적용과정은 당일 방문 고객  $u$ 에 대한 맞춤형 카드를 추천하기 위해 학습과정에서 도출한 IUN을 이용하여 해당 고객에게 적합한 신용카드를 실시간으로 추출하는 과정이다. 실제 환경에 적용할 때 <Figure 1>의 학습과정은 고객프로파일의 변화에 따라 주기적으로 반복 수행되어 IUN을 갱신하여야 하며, 적용 과정은 이렇게 갱신된 IUN을 기반으로 작동된다.

적용과정은 <Figure 1>에서와 같이 임의의 고객  $u$ 가 방문했을 때 시작되며, 먼저 학습과정에서 도출한 IUN을 기반으로 고객  $u$ 와 유사한  $k$ 명의 이웃 고객을 선택한다. 이때 유사도를 계산하는 방식은 학습과정과 동일하다.

다음은 고객  $u$ 를 위한 맞춤형 추천 카드를 도출하는 과정으로,  $u$ 의  $k$ 명의 이웃 고객들이 보유한 카드 중에서 고객  $u$ 가 보유하지 않은 카드를 추천 대상 카드로 선택한다. 만일 이웃 고객들이 보유한 카드의 종류가  $n$ 개 보다 많을 경우에는 가장 많이 발급된 카드 순으로  $n$ 개를 선택한다.

이렇게 선택된 추천카드는 영업점을 방문한 고객에게 맞춤 카드로 추천되거나, 그 외에도 홈페이지, 스마트폰의 앱, SMS, TM, DM 등 다양한 광고수단을 통해서 추천될 수 있다.

## 3. 실험

### 3.1 실험 데이터

본 연구에서는 제안된 기법의 성능 및 적용가능성을 검증하기 위하여 국내 A은행의 신용카드 고객을 대상으로 실험을 진행하였다.

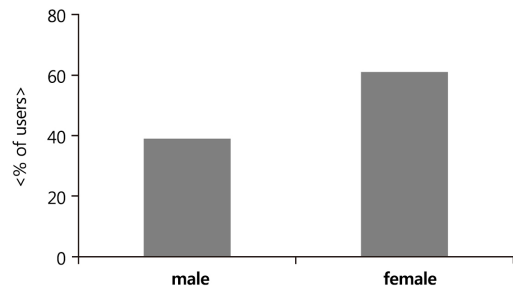
전체 고객 중 3,564명의 고객을 표본 추출하였고, 고객프로파일을 구성하기 위하여 실험 고객의 인구통계학 정보와 1년 동안의 소비패턴, 거래수단, 보유상품, 가입 서비스 정보를 활용하였다. 한편, A은행에서 취급하는 신용카드 수는 152종이다.

실험에 활용한 데이터의 특성은 <Table 2>에 정리되어 있다.

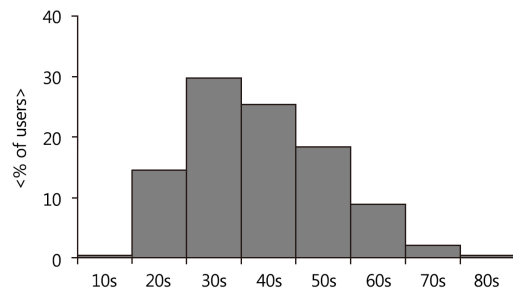
<Table 2> Users Data

period	2016. 1~2016. 12
no. of users	3,564
no. of credit & check card	152

실험고객의 인구통계학적 특성 중 성별과 연령 분포는 <Figure 7>, <Figure 8>에 각각 표시되어 있다. 여성이 61%, 남성이 39%로 여성 고객이 많았으며(<Figure 7>참조), 연령별로는 30대가 29%, 40대가 25%, 50대가 18%로 전체 고객의 72%를 차지하였다(<Figure 8> 참조).

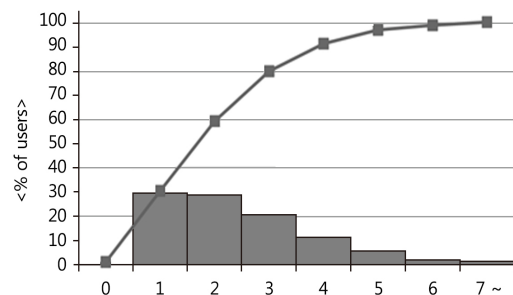


<Figure 7> Gender Ratio



<Figure 8> Age Distribution

한편, <Figure 9>에서 고객이 보유한 카드수의 분포를 살펴보면 3개 이하의 카드를 보유한 고객이 2,838명(80%)으로 대부분을 차지하고 있음을 알 수 있다.



<Figure 9> Distribution of the Number of Possessed Card

### 3.2 분석결과

최근 관련 연구에서 많이 사용하는 F1 값을 이용하여 제안된 기법의 성능을 분석하였다. 기존의 많은 연구들이 추천 시스템의 성능을 평가하기 위하여 정확도, 재현율 및 F1 값을 사용해 왔다[1, 24, 26]. 정확도는 추천시스템이 추천한 아이템 중 실제 고객이 구매하는 아이템의 비율

이며, 재현율은 고객이 구매한 아이TEM 중에서 추천 시스템의 추천이 적중시킨 비율이다.

한편, 정확도와 재현율은 시스템이 추천하는 아이TEM의 개수  $n$ 에 영향을 받는다. 즉,  $n$ 이 많아지면 재현율은 올라가고 정확도는 떨어지게 된다. 따라서 본 연구에서는 정확도와 재현율을 동시에 고려할 수 있는 F1 값을 사용하여 추천시스템의 성능을 분석한다.

실험을 위해 A은행의 신용카드 가입 고객 중에서 표본 추출한 3,564명의 데이터를 대상으로 학습과정을 진행하였다. 제안된 방법론에 따라 5개의 고객프로파일 및 고객 네트워크를 도출하였고 이를 기반으로  $IUN$ 을 구성하였다. 다음으로 성능의 분석을 위해 전체 데이터의 30%를 테스트 데이터로 구성하였으며,  $IUN$ 을 기반으로 테스트 데이터에 포함된 고객을 대상으로 카드 추천을 수행하였다. 이후 해당 고객들이 실제 보유한 카드와 추천된 카드의 비교를 통해 그 성능을 분석하였다.

본 실험에서는 제안된 기법을 대상으로 정확도, 재현율 및 F1 값을 산출하여 그 성능을 살펴볼 것이며, 또한 성능의 비교평가를 위하여 베스트셀러 기반 추천기법과 그 값을 비교해 본다.

3.2.1 이웃 및 추천카드 수의 변화에 따른 성능 비교

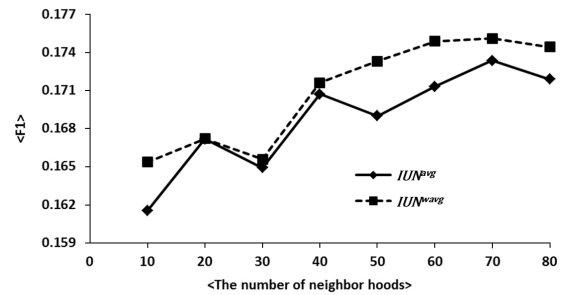
먼저 제안된 기법의 성능을 확인하기 위해 이웃의 수  $k$  및 추천 카드의 수  $n$ 을 변화시키면서 성능을 평가하였다. 다음과 같이 5종의 개별 고객 네트워크와 2종의 통합 고객 네트워크를 구성하여 제안된 알고리즘의 성능을 확인하였다.

- $UN_c$  기반 추천 : 단일 프로파일 방식으로 사용자의 소비성향을 알 수 있는 소비패턴 프로파일만을 사용하여 추천하는 방식
- $UN_g$  기반 추천 : 단일 프로파일 방식으로 사용자의 금융상품 가입성향을 알 수 있는 보유상품 프로파일만을 사용하여 추천하는 방식
- $UN_d$  기반 추천 : 단일 프로파일 방식으로 인구통계학 프로파일만을 사용하여 추천하는 방식
- $UN_i$  기반 추천 : 단일 프로파일 방식으로 사용자의 금융업무 이용수단을 나타내는 거래수단 프로파일만을 사용하여 추천하는 방식
- $UN_s$  기반 추천 : 단일 프로파일 방식으로 사용자의 비대면 금융업무 선호도를 나타내는 가입서비스 프로파일만을 사용하여 추천하는 방식
- $IUN^{avg}$  기반 추천 : 다중 프로파일 방식으로  $IUN$  구성시, 5종의 개별네트워크의 유사도를 단순 평균방식으로 결합하는 방식
- $IUN^{wavg}$  기반 추천 : 다중 프로파일 방식으로  $IUN$ 을

구성 시, 5종의 개별네트워크의 유사도를 각 네트워크의 F1 값을 가중치로 가중 평균하여 결합하는 방식

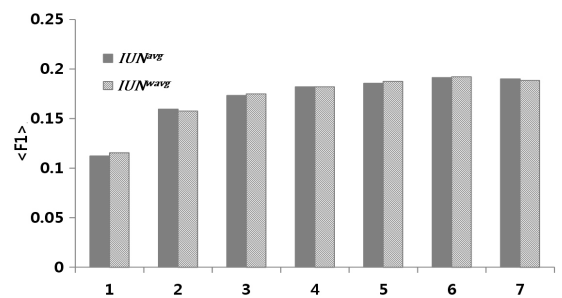
<Figure 10>은 통합네트워크 구축 방법에 따른 제안된 기법의 성능을 보여주고 있다. 현실적으로 영업점 방문 고객에게 추천할 수 있는 카드 수의 최대는 3개 이하라는 현장 실무자의 조언에 따라, 추천 카드 수  $n$ 을 3으로 하고 이웃수  $k$ 를 10에서 80까지 10명씩 증가시켜 가면서 F1 값의 변화를 나타내었다.

<Figure 10>에 따르면  $IUN^{wavg}$  기법이  $IUN^{avg}$  보다 항상 좋은 성능을 보이고 있으며, 이웃수  $k$ 가 70일 때  $IUN^{wavg}$ 의 F1 값이 0.175로 가장 높음을 알 수 있다. 이는  $IUN$  구성시 5종 프로파일의 각 네트워크가 추천 성능에 기여하는 정도를 가중치로 반영하는 방법이 보다 뛰어난 성능을 보임을 의미한다.



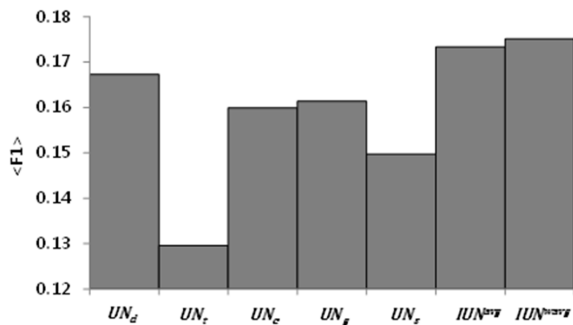
<Figure 10> F1 Value of Ensemble  $IUN$  for Various  $k$

<Figure 11>은 이웃수  $k$ 를 70으로 고정했을 때, 추천 카드의 수  $n$ 의 변화에 따른 제안된 기법의 성능을 보여주고 있다. <Figure 11>에 따르면, 추천카드의 수가 증가할수록 제안된 기법의 F1 값은 상승하다가  $n=6$ 에서  $IUN^{wavg}$ 의 F1 값이 0.192로 최댓값을 갖고 이후 감소하는 경향을 보인다. 이는 F1 값을 기준으로 추천 카드의 수를 6개로 했을 때 가장 높은 성능을 보임을 의미하지만, 앞서 언급한 것처럼 3개 이하의 카드를 추천하는 것이 현실적이므로, 이후의 비교에서는 추천카드의 수를 3으로 고정한다.



<Figure 11> F1 Value for Various  $n$

<Figure 12>는 단일 프로파일 기반 추천방식과 다중 프로파일 기반 추천방식의 성능을 보여주고 있다. 이 실험에서는  $k=70$ ,  $n=3$ 을 사용하였다. <Figure 12>에 따르면 단일 프로파일 기반의 추천 방식 보다 다중 프로파일을 통합한 추천 기법의 성능이 더 우수함을 알 수 있다. 단일 프로파일 기반 추천 방식 중에는 인구통계학 프로파일을 기반으로 한 추천 방식이 가장 높은 성능을 보이고 있으나,  $IUN^{wavg}$ 가 이에 비해 약 4.6% 더 높은 성능을 보이고 있다.

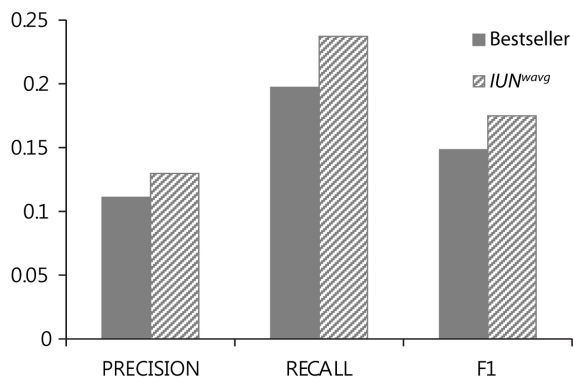


<Figure 12> Performance Comparison of Various UNs

### 3.2.2 다른 기법과의 성능비교

다음으로 제안된 기법의 객관적인 성능비교를 위하여 실험기간 동안 가장 가입 빈도가 높은 신용카드를 모든 고객에게 동일하게 추천하는 방식인 베스트셀러(bestseller) 기반 추천[21]과 제안된 기법을 비교하였다. 베스트셀러 기반 추천기법은 개별 고객의 고객프로파일에 대한 고려 없이, 학습기간 동안 가장 발급빈도가 높은 신용카드를 모든 고객에게 동일하게 추천하는 방식이다.

본 비교를 위해 사용된 제안 기법은 가중평균으로 통합한  $IUN^{wavg}$  기법이며, 추천카드의 수  $n$ 은 3개, 이웃 수 70을 사용하였고, 베스트셀러 기법에서도 가장 많이 발급된 순서대로 3개의 카드를 추천하였다.



<Figure 13> Performance Comparison of Proposed Scheme and Bestseller Scheme

<Figure 13>에 따르면 제안된 기법의 정확도, 재현율, 그리고  $F1$  값은 베스트셀러 기반 추천방식보다 높음을 알 수 있다. 베스트셀러 기반 추천의  $F1$  값이 0.148이며 제안된 기법은 0.175로, 베스트셀러 기반 추천에 비해서 18.0% 우수한 성능을 보이는 것으로 확인되었다.

## 4. 결 론

본 연구에서는 고객 맞춤형 신용카드 추천을 위하여 콜드스타트 문제를 해결할 수 있는 다중 프로파일 기반의 협업필터링기법을 제안하였다. 제안된 기법은 학습과정과 적용과정으로 구성된다. 학습과정에서는 사용자 정보를 기반으로 5종의 다중 프로파일을 구성하고, 이로부터 통합 고객 네트워크를 도출한다. 다음으로 적용과정에서는 학습과정에서 구축된 통합 고객 네트워크를 활용하여 유사고객을 선별하고 유사고객이 보유한 신용카드를 방문 고객에게 추천한다.

본 연구에서는 제안된 기법의 성능을 평가하기 위하여 국내 소재 A은행의 신용카드 자료를 확보하여 실험을 수행하였다. 실험결과 제안된 다중 프로파일 기반의 추천기법은 단일 프로파일 및 베스트셀러 기반 추천기법과 비교하여 높은 추천 정확도를 보였다.

본 연구는 신용카드 추천 과정에서 발생할 수 있는 콜드스타트 문제를 해결할 수 있다는 것에 큰 의미가 있다. 신용카드는 온라인 쇼핑과는 달리 보유 상품의 수가 제한적이고 그 수요가 지속적·반복적으로 발생하지 않는다는 특징을 갖는다. 따라서 일반적인 추천시스템과 같이 사용자의 상품 선호도에 기반을 둔 방식을 적용할 때는 고객 정보의 부족으로 콜드스타트 문제가 발생하게 되고, 이는 추천시스템의 성능 저하를 초래하게 된다. 이러한 문제를 해결하기 위해 본 연구에서는 고객의 신용카드 선호도를 제외한 다양한 고객 정보로부터 다중 프로파일을 구성하여 카드 추천에 활용하였다. 또한, 본 연구의 결과는 다양한 상품을 취급하는 금융기업이 고객의 특성에 맞는 금융상품을 효과적으로 추천하여 고객의 만족을 높이고 기업 수익 증대하는데 활용할 수 있을 것으로 기대된다.

한편, 본 연구의 한계점 및 향후 연구과제는 다음과 같다.

첫째, 본 연구에서 제안한 다중 프로파일 기반의 추천기법은 고객프로파일이 등록되지 않은 고객에게는 적용할 수 없다는 한계를 갖고 있다. 고객정보가 등록되어 있지 않은 신규고객은 창구 담당직원 상담, 베스트셀러 기법 등의 추천방식이 더욱 적합할 것이다.

둘째, 본 연구에서는 고려하지 않았지만 실제 적용 환경



에서 학습데이터를 구성할 때는 프로모션 등의 특수한 상황에서 발생한 상품 가입 및 고객 정보는 배제하는 것이 필요하다. 이러한 데이터가 학습데이터에 포함되는 경우 학습 결과를 왜곡하고 추천 성능을 저해하는 결과를 초래할 수 있다.

셋째, 실제 서비스 환경에서 각 고객의 신용카드 구매 패턴은 현장 직원의 맞춤형 추천에 의해 영향을 받을 수 있으나, 이전의 다수의 연구에서와 마찬가지로 사후적인 고객 구매행위 데이터를 통한 실험만으로는 이러한 효과를 파악하는데 한계가 존재한다.

추후 연구에서는 신용카드 이외의 주요 금융상품인 방카슈랑스(bancassurance)나 펀드 등을 대상으로 맞춤형 추천기법을 제안하고자 한다.

## Acknowledge

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2015S1A5A8012533).

## References

- [1] Ahn, H.J., A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, *Information Sciences*, 2008, Vol. 1, No. 78, pp. 37-51.
- [2] Ali, K. and Stam, W.V., Tivo : Making show recommendations using a distributed collaborative filtering architecture, *Proceedings of the Conference on Knowledge Discovery and Data Mining*, Washington, USA, 2004, pp. 394-401.
- [3] Bell, R., Koren, Y., and Volinsky, C., The Bellkor 2008 Solution to the Netflix Prize, 2008.
- [4] Bennet, J. and Lanning, S., The Netflix Prize, *Proceedings of the KDD CUP and Workshop*, 2007.
- [5] Cho, Y.H., Park, S.K., Ahn, D.H., and Kim, J.K., Collaborative Recommendations using Adjusted Product Hierarchy : Methodology and Evaluation, *Journal of the Korean Operations Research and Management Science Society*, 2004, Vol. 29, No. 2, pp. 61-77.
- [6] Choi, J.S., Kim, D.Y., Han, I.S., and Kim, Y.C., A Study on Data Analysis in on/off-line Shopping mall, *Electronic Commerce Research Association*, 2011, Vol. 12, No. 2, pp. 71-88.
- [7] Credit FINANCE Association, <http://www.crefia.or.kr>.
- [8] Das, A., Datar, M., Garg, A., and Rajaram, S., Google news Personalization : Scalable online Collaborative filtering, *Proceedings of the World Wide Web Conference*, 2003, Budapest, Hungary, pp. 271-280.
- [9] Financial Supervisory Service, <http://www.fss.or.kr>.
- [10] Jeong, S.B., The Relationship between Structure Attributes of Product Networks and Performance of integrated Internet Shopping Malls, *Association of Business Education*, 2015, Vol. 30, No. 3, pp. 97-106.
- [11] Kim, C.-J. and Choi, K.-S., A Mobile Web's Recommendation Technique based on XPDL, *Journal of the Korea Academia-Industrial Cooperation Society*, 2013, Vol. 14, No. 11, pp. 5856-5865.
- [12] Kim, K.H. and Lee, B.W., A Marketing Study of Credit Card with Aging Society, *The Credit Card Review*, 2012, Vol. 6, No. 1, pp. 87-105.
- [13] Kim, K.-J. and Ahn, H., Collaborative Filtering with a User-Item Matrix Reduction Technique for Recommender Systems, *International Journal of Electronic Commerce*, 2011, Vol. 16, No. 1, pp. 107-128.
- [14] Kim, K.-J. and Ahn, H., User-Item Matrix Reduction Technique for Personalized Recommender Systems, *Journal of Information Technology Applications & Management*, 2009, Vol. 16, No. 1, pp. 97-113.
- [15] Kim, M. and Cho, Y., A Multimodal Profile Ensemble Approach to Development of Recommender System Using Big Data, *The Journal of Intelligence and Information System*, 2015, Vol. 21, No. 4, pp. 93-100.
- [16] Lee, J.S. and Park, S.D., Performance Improvement of a Movie Recommendation System using Genre-wise Collaborative Filtering, *Journal of Intelligence and Information System*, 2007, Vol. 13, No. 4, pp. 65-78.
- [17] Lim, C.-H., On-line Recommendation Service Algorithm using Human Sensibility Ergonomics, *Journal of the Society of Korea Industrial and Systems Industrial and Systems Engineering*, Vol. 27, No. 1, pp 38-46.
- [18] Linden, G., Smith, B., and York, J., Amazon.com recommendations : Item-to-item Collaborative filtering, *IEEE Internet Computing*, 2003, Vol. 7, No. 1, pp. 76-80.
- [19] Nam, K.W., Park, S.S., Shin, Y.K., and Jang, D.S., An Analytical Study on Research trends of Web-based Recommendation System, *Proceedings of the Korean Operations Research and Management Science Society Spring Joint Meeting*, Pohang, Korea, 2008, pp. 86-97.
- [20] Niwattanakul, S., Singthongchai, J., Naenudorn, E., and Wanapu, S., Using of Jaccard Coefficient for Keywords Similarity, *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, 2013,

Vol. 1, pp. 13-15.

- [21] Park, J.H., Cho, Y.H., and Kim, J.K., Social Network : A Novel Approach to New Customer Recommendations, *Journal of Intelligent Information Systems*, 2009, Vol. 15, No. 1, pp. 123-140.
- [22] Park, S. and Pennock, D., Applying Collaborative filtering techniques to movie search for better ranking and browsing, *Proceedings of 13<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, California, USA, 2007, pp. 550-559.
- [23] Piotte, M. and Chabbert, M., The Pragmatic theory solution to the Netflix grand prize, Netflix prize documentation, 2009.
- [24] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J., Analysis of Recommendation Algorithm for E-Commerce, *Proceedings of the ACME-commerce conference*, Minnesota, USA, 2000, pp. 158-167.
- [25] Schein, A.I., Popescul, A., Ungarnd, L.H., and Pennock, D.M., Methods and Metrics for Cold-Start Recommendations, *Proceedings of the 25<sup>th</sup> Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tampere, Finland, 2002, pp. 253-260.
- [26] Schonberger, V.M. and Cukier, K., Big data : a revolution that will transform how we live work and think, John Murray, 2013.

#### ORCID

Won Cheol Lee | <http://orcid.org/0000-0002-5529-518X>

Hyoup Sang Yoon | <http://orcid.org/0000-0003-3306-7327>

Seok Bong Jeong | <http://orcid.org/0000-0002-6209-1935>