

## XML 기반 여러 형태 프로파일을 이용한 고객세분화

안형근<sup>o</sup>, 이단영, 고재진  
울산대학교 컴퓨터·정보통신공학부  
{hkahn<sup>o</sup>, danyoung64, jkkoh}@mail.ulsan.ac.kr

### Client Segmentation using XML-based Multiform Profile

Hyoung-Keun An<sup>o</sup>, Dan-Young Lee, Jae-Jin Koh  
School of Computer Engineering & Information Technology, University of Ulsan

#### 요약

최근 정보 통신기술의 발전으로 인하여 전자상거래가 확산되고 있는 실정이며, 이용하는 고객 또한 상당히 증가하고 있다. 고객의 활발한 구매 거래 활동으로 하루에도 아주 많은 양의 데이터가 생성되고 있는 실정이다. 이에 전자상거래의 웹 사이트 관리자나 경영자는 고객의 구매형태나 패턴의 특징을 파악하여 보다 효율적인 서비스를 고객에게 제공하기 위하여 현재까지 유사그룹의 고객 세분화를 적용하는 연구가 이루어지고 있다. 본 논문에서는 전자상거래에서 고객들의 정보를 분석하여 개인화하기 위한 방법으로 사용되는 고객 프로파일을 이용하여 고객세분화 하는데 적용을 하고자 한다. 기존 고객세분화의 통계적인 분석이 아닌 XML 기반의 고객 정보를 XPath를 이용하여 고객세분화에 필요한 규칙을 생성하고, 그 규칙을 바탕으로 고객 프로파일을 생성하는 방법과 프로파일을 이용한 군집에 따른 분석 결과 및 추천서비스를 소개하고자 한다.

## 1. 서론

정보 통신기술인 네트워크 기술의 발전과 인터넷 기술의 발달로 인하여 여러 형태의 웹 서비스가 제공되고 있으며 그 중에서 가장 많은 관심을 가지는 것은 전자상거래일 것이다. 전자상거래의 장점으로는 공간적, 시간적 구매 자유와 다양한 상품을 구매할 수 있다는 사실은 누구나 인지를 하고 있을 것이다. 이와 같은 장점을 극대화하기 위해서는 전자상거래 상에서의 많은 상품들 중에 고객이 원하는 상품을 효과적으로 어떻게 제공할 것인지, 전자상거래 운영자는 미래에 있어 고객이 어떤 상품을 원하는지를 예측할 수 있는 연구와 방법들이 현재까지 많이 이루어지고 있다. 최근까지 데이터마이닝이나 CRM시스템 등을 이용하여 고객을 유지, 관리 및 서비스를 하고 있지만, 이러한 시스템을 개발하고 관리하는데 있어서는 과도한 비용과 시간이 소요되는 점과 관리자가 시스템을 올바르게 이해하고 관리하는데 있어 많은 어려움이 있다.[1] 이를 보완하기 위해 좀 더 쉽고 이용이 간편하며 비용이 적게 드는 개인화 서비스에 대한 연구가 활발하게 진행되고 있다.

따라서 본 논문에서는 개인화 추천에 필요한 사용자 프로파일을 이용하여 전자상거래에서 빠르게 고객을 분류하여 분석할 수 있는 고객세분화 방법에 적용하고자 한다. 해당 논문에서 제시하는 사용자 프로파일은 구조적인 XML기반의 고객정보와 DTD(Data Type Definition)를 이용하여 DOM(Document Object Model) 트리를 구성하며, XPath를 이용하여 프로파일을 생성할 수 있는 규칙 집의 리스트를 만들어 저장한다. 다양한 규칙 리스트는 고객 프로파일을 생성시키고 그리고 고객세분화에 적용이 될 수 있다.[5] 여러 형태의 규칙에 의한 고객 프로파일 생성으로 인하여 세분화된 다양한 고객 분석에도 도움을 줄 수 있을 것이다. 또한, 개인화 상품 추천서비스도 다양하게 지원해 줄 것이다.

## 2. 관련 연구

### 2.1 CRM과 eCRM

CRM(Customer & Relationship Management)이란 고객들에게 더 좋은 서비스를 제공하고 고객들과의 상호작용을 효율적으로 관리함으로써 기업이 그들의 가장 바람직한 고객에게 좀 더 효과적인 차별화를 제시하고, 자원을 효과적으로 배치하도록 하여 기업의 수익을 증대시키기 위한 전략이다.

eCRM(electronic Customer & Relationship Management)이란 인터넷에서 상품을 구매하는 고객이나 서비스를 구매하는 고객을 대상으로 인터넷 기반 기술을 이용하여 신규 고객의 획득, 기존 고객의 유지 등을 통하여 기업의 수익성을 최적화 하고자 하는 전략이다.

### 2.2 개인화(Personalization)

개인화 전략은 각 사용자별로 맞춤 서비스를 제공하기 위한 핵심 전략중의 하나이다. 콘텐츠를 제공하는 사이트에서 각 사용자의 개인화를 가능하게 하기 위해서 사용자 프로파일을 이용한다. 이러한 각 사용자마다의 프로파일을 이용함으로써 각 사용자들에게 차별화된 서비스를 제공해 고객들과의 친밀한 관계를 유지하여 충성도를 높일 수 있다. 또한 콘텐츠 사이트에서 개인화 전략은 사용자들의 콘텐츠를 검색하는 시간과 경비를 절약하고 사이트에 대한 만족도를 높여 매출을 극대화하기 위한 전략으로 사용할 수 있다.

### 2.3 고객세분화(Client Segmentation)

고객세분화는 많은 고객들을 서로 비슷한 소비자들끼리 분류하는 작업을 말한다. 고객의 성별 또는 나이, 직업, 지역 등에 따라 나누는 것도 세분화 작업이다. 세분화는 전체 데이터베이스를 바라볼 수 있도록 해주거나, 기업이 서로 다른 등급에 맞는 적절한 마케팅이 가능하게 해준다. 데이터 기반의 세분화는 다양한 통계적 기법과 데이터마이닝을 사용하여 수행되어 지는데 대부분은 의사결정 나무(Decision Tree)와 군집분석(Cluster Analysis) 그리고 신경망(Neural Network)을 이용한 기법 등 세 개로 분류된다. 본 논문에서는 전자상거래의 고객층이 다양

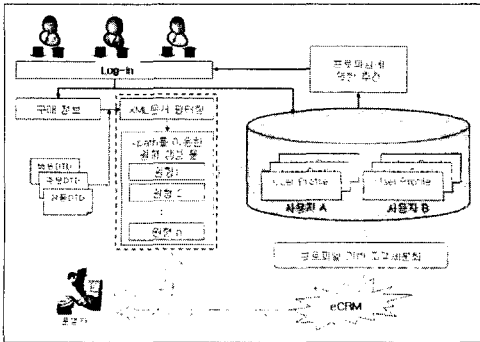
하고 많으므로 개인 프로파일을 고객세분화의 군집분석 방법에 적용하고자 한다.

### 2.4 XPath

XML은 문서의 요소들을 노드(node) 개념으로 접근하여 소스 트리의 정확한 위치를 지정해주기 위한 경로지정 문법으로 XML과 별개의 문법체계를 가지고 있다. XPath 단독으로 사용하지 않으며, XSL, XSLT에서 결과 트리를 만들기 위한 경로 지정에 사용된다. 또한 XPath는 XSLT, XLink, XPointer에 의해 사용될 수 있는 XML 문서의 내용을 질의하고 위치를 부여하기 위한 일반 문법과 의미를 제공하기 위해 만들어진 언어이다.

## 3. 프로파일을 이용한 고객세분화 시스템

본 논문에서 제안하고자 하는 것은 특정 전자상거래에서 관리자가 필요로 하는 고객과 분석에 필요한 여러 형태의 규칙, 프로파일을 생성하여 eCRM에 필요한 분석 정보를 추출하고, 그 분석 결과로 세분화에 따른 고객들에 적당한 정보를 추천할 수 있는 시스템을 설계 및 구현하고자 한다. (그림 1)은 본 논문에서 제시하는 시스템의 전체적인 흐름을 보여주고 있다.



(그림 1) 전체 시스템 흐름

### 3.1 프로파일을 위한 규칙 생성

고객을 분석하기 위한 관리자가 정의하는 규칙은 다음과 같이 정의 할 수가 있다.

$$PT_n = \{PT_{desc} \langle select\_elements \rangle\} \quad n=1,2,\dots,n \quad (식 1)$$

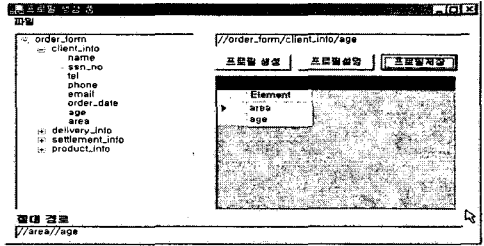
$PT_n$ 는 프로파일 생성을 위한 규칙이며, 규칙은 다수 개 정의 할 수가 있다.  $PT_{desc}$ 는 생성규칙의 설명을 말하며,  $select\_elements$ 는 프로파일 생성을 위한 요소이다. 예로, 고객의 거주 지역과 나이를 대상으로 고객을 분석하고자 한다면 다음과 같이 규칙을 정의 할 수가 있다.

$$PT_1 = \{PT_{지역과나이정보} \langle area, age \rangle\}$$

이 프로파일 규칙은 다음과 같이 분석될 수가 있을 것이다. 같은 지역 및 비슷한 나이 대에 따른 고객들의 정보를 필터링 하여 수집하고 고객을 세분화하여 관리자는 세분화에 따른 군집의 성향을 분석하고, 그 결과를 고객들에게 적당한 정보를

추천하는 프로파일 생성을 위한 규칙이라 할 수 있다.

본 논문에서는 거래되는 모든 문서가 XML 기반을 기반으로 한다고 본다. 따라서 XML의 구조적인 정보를 필터링하기 위한 방법으로 DTD를 이용하여 DOM 트리구조로 표시하여, 프로파일을 생성에 필요한 항목을 선택, 생성할 수 있는 GUI 툴을 구현하였다. 선택된 항목들은 XML 문서를 검색 필터링하기 위하여 XPath 형태로 저장이 된다. 아래 (그림 2)는 규칙을 생성하기 위한 XPath를 이용한 규칙 생성 GUI 툴을 보여주고 있다.



(그림 2) DTD를 이용한 XPath기반 규칙생성 툴

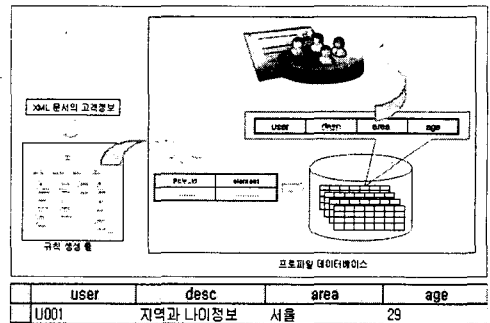
### 3.2 규칙에 따른 사용자 프로파일 생성

전자상거래에서 고객이 로그인한 후 기존 고객 및 신규 고객의 트랜잭션(거래) 후의 XML 문서는 필터링하여 생성된 규칙에 의해서 여러 형태의 프로파일을 생성한다. 앞 3.1절에서 규칙을 지역(area)과 나이(age)로 설정을 하였다. 이들 규칙은 XML 문서의 정보 추출을 위하여 XPath 형태로 저장 된다. 지역과 연령의 XPath 표현은 다음과 같이 나타낸다.

- 지역정보 : //order\_form/client\_info/area
- 나이정보 : //order\_form/client\_info/age

필터링 과정에서 규칙과 관련된 요소들만 XML 문서에서 추출된다. 지역과 나이는 XML 문서에서 <area>, <age>와 같은 XML 태그로 정의되어 있다. 다음은 규칙에 의해 생성된 ID : u001의 프로파일을 보여주고 있다.

$P_{u001} = \{PT_{지역과나이정보} \langle area \rangle, \langle age \rangle\}$ . u001의 고객이 다음 로그인시에는 해당 프로파일에 의한 다양한 추천서비스를 받을 수가 있다.

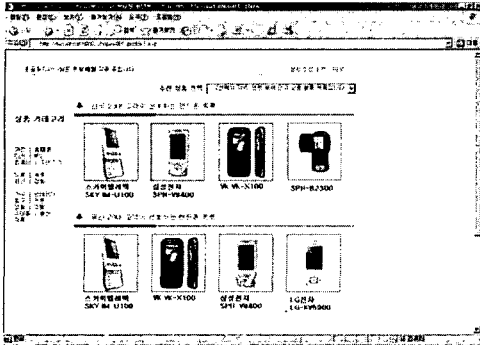


(그림 3) 프로파일 생성 흐름 및 데이터베이스

### 3.3 고객 세분화와 프로파일에 의한 추천서비스

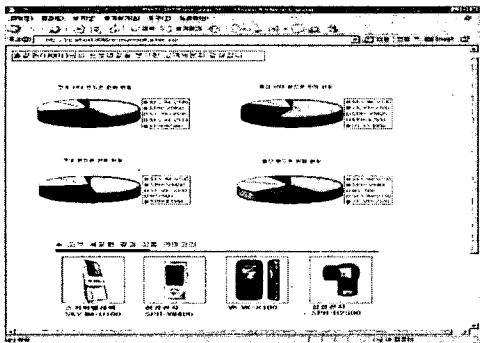
규칙 생성 툴에 의해서 생성된 사용자 프로파일을 기반으로 고객이 로그인을 하게 되면 프로파일과 유사한 그룹을 적용한 추천 화면을 보여주게 된다. (그림 4)는 고객의 지역과 나이 프로파일을 바탕으로 추천한 핸드폰 상품 목록을 보여주고

있다. 또한 상단에 프로파일에 의한 분류별 선택할 수 있는 메뉴가 있어 다양한 추천 서비스 상품들을 볼 수가 있다.



(그림 4) 고객 세분화에 따른 상품 추천화면

(그림 5)는 전자상거래 운영자가 규칙에 의해 생성된 프로파일을 대상으로 고객을 세분화한 결과 분석 화면을 보여 주고 있다. 이 분석을 통하여 고객들의 성향을 파악할 수가 있어 eCRM에 많은 도움을 줄 수가 있을 것이다.



(그림 5) 프로파일에 의한 고객세분화 분석화면

#### 4. 결 론

전자상거래에서는 고객들의 선호도 파악을 위해서 여러 가지의 분석 기법을 사용하고 있다. 이러한 기법을 수행하기 전에 데이터를 적절한 형태로 가공하기 위한 전처리 작업이 필요하며, 이 목적으로 운영되는 시스템에 맞는 규칙을 생성하여 유지하기 위해서는 많은 비용이 든다.

본 논문에서는 전자상거래에서 효율적인 서비스를 제공하고 자하는 시스템에서 활용될 수 있는 고객 프로파일과 규칙 생성에 필요한 XPath 기반의 GUI 툴을 제시하였다. 해당 툴은 XML 기반의 고객정보를 DTD에 의해서 DOM 트리로 구성한 다음 프로파일 생성에 필요한 규칙 항목을 선택할 수가 있었다. 생성된 규칙에 의해서 다양한 고객 프로파일을 생성 시킬 수가 있다. 전자상거래는 고객층이 다양하기 때문에 빠른 분석을 위해서 고객세분화 기법에 프로파일을 적용을 하였다. 결과 전자상거래 운영자는 분석을 위한 그룹의 규칙을 생성하고 고객들의 정보를 필터링하여 해당 그룹의 성향을 즉시 분석할 수가 있으면, 고객은 로그인 후 유사한 그룹에 따른 다양한 추천 서비스를 받을 수가 있다.

천 서비스를 받을 수가 있다.

본 논문의 여러 형태의 프로파일을 이용한 고객 세분화의 방법으로 eCRM(고객관리업무), 개인화 전략, 프로파일을 필요로 하는 마케팅 또는 영업활동 분야에 응용 기대된다.

#### 5. 참고 문헌

- [1] G. Adomavicius and A.Tuzhilin, Using Data Mining Methods to Build Customer Profiles. IEEE, February, pp74~82, 2001.
- [2] Ugur Cetintemel, Michael J. Franklin, and C. Lee Gils. Self-Adaptive User Profiles for Large-Scale Data Delivery, Proceedings of the International Conference on Data Engineering, San Die해, CA, pp 622~633, February, 2000.
- [3] D. Aksoy, M. Altinel, etc., "Research in Data Broadcast and Dissemination", Proc. 1st Intl. Conf. on Advanced Multimedia Content Processing, Osaka, Japan, November, 1998.
- [4] T. W. Yan, H. Garcia-Molina, "Index Structures for Selective Dissemination of Information Under Boolean Model", ACM TODS, 19(2):332~364, 1994.
- [5] Mehmet Altinel, Michael J. Franklin, "Effective Filtering of XML Document for Selective Dissemination of Information", Proceedings of the 26th VLDB Conference, Cairo, Egypt, 2000.