

# 전자상거래 소비자 구매 패턴 분석 도구 개발

한지선<sup>o</sup> 조동섭  
이화여자대학교 컴퓨터학과  
{nabi77, dscho}@ewha.ac.kr

## Design and Implementation of Customer's Buying Trend Analysis in e-Commerce Environment

Ji-Seon Han<sup>o</sup> Dong-Sub Cho  
Dept. of Computer Science and Engineering, Ewha Womans University

### 요 약

전자상거래에서는 소비자의 구매 패턴을 분석하는 것이 필요하다. 이러한 패턴을 효과적으로 분석하기 위해 지능형 로그 서버를 정의하고 이를 설계, 구현하였다. 지능형 로그 서버란 전자상거래 쇼핑물의 환경에서 사용자 로그를 데이터베이스화하여 저장하고 데이터베이스에 저장된 정보를 서버 종류별, 시간별, 페이지별 등으로 분석하여 사용자 패턴을 분석할 수 있는 서버를 말한다. 이 서버는 텍스트 파일로 로그를 저장하는 서버보다 자세한 정보를 효율적으로 저장할 수 있다. 그리고 데이터베이스 접근 기술로 ADO(ActiveX Data Object)를 사용하여 데이터베이스 접근 속도를 향상시켰으며 관계형·비관계형 데이터베이스에 모두 접근할 수 있다는 장점을 가진다. 또한 소비자의 구매 패턴을 분석하기 위해 DBMiner2.0을 사용하였다.

### 1. 서론

인터넷을 배경으로 한 전자상거래의 등장은 전통적인 무역거래의 관념과 패턴을 뒤바꾸어 놓았으며 시간과 공간의 제약을 훌쩍 뛰어넘어 언제 어디서나 원하는 상품의 거래가 가능하도록 하였다[1].

전자상거래는 단순히 인터넷상에서 물건을 사고 파는 수준을 넘어 제조·유통, 금융산업은 물론 언론, 교육, 행정 등 사회 전 분야에 영향을 미치고 있다.

개인이나 조직, 기업은 인터넷을 이용하여 이러한 전자상거래 시스템에 접근한다. 그러므로 이 시스템은 쌍방향 커뮤니케이션을 바탕으로 운영된다. 따라서 전자상거래 시스템의 원활한 운영을 위해서 인터넷이란 강력한 매체를 이용하여 고객 지원, 글로벌 전자 마케팅, 온라인 상품 전시 등의 마케팅을 위한 다양한 서비스를 지원할 수 있어야 한다[2]. 그리고 이러한 서비스 지원을 위해 소비자의 구매 패턴 분석이 필요한 것이다.

그러나 현재의 전자상거래 사이트들은 소비자들이 어디서, 어떻게, 왜 방문하는지 등에 대한 분석이 미비하다[3]. 게다가 쇼핑물의 경우 어떤 소비자들이, 언제 구매 의욕이 생기는지, 어떻게 어느 정도의 시간이 걸려서 구매를 하게 되는지 등의 구매 패턴에 대한 분석은 더욱 미비한 실정이다.

그러므로 본 논문에서는 위와 같은 서비스를 소비자의 성향에 맞게 제공하기 위해 데이터 분석을 수행하는 지능형 로그 서버를 제안하고 구현하고자 한다.

### 2. 관련연구

#### 2.1 기존의 전자상거래 시스템

기존의 전자상거래 시스템에서 웹 로그 분석을 위한 시스템은 웹서버와 웹 로그 분석 서버로 구성된다. 웹서버는 웹 로그를 생성하고 웹 로그 분석 서버는 생성된 로그를 분석하는 작업을 수행한다.

그 수행 과정을 살펴보면 다음과 같다.

step1 : 웹서버는 접근, 예러, 참조, 에이전트 등에 대한 로그를 파일로 저장

step2 : 웹 로그 분석 서버는 저장된 로그 파일에 대해 분석

step3 : 분석 결과를 가시화해서 웹페이지 형식으로 저장

step4 : 사용자는 웹을 이용하여 확인

이상의 수행과정에서 다중의 로그 파일이 생성되고 각 로그 파일에 대해 분석하는 오버헤드가 발생한다. 그리고 로그 파일에는 IP주소나 사용브라우저, 날짜 같은

정보를 정확히 구분하지 않고 공백이나 탭 등을 이용하여 저장하므로 분석 시 모호한 부분이 발생할 수 있다.

뿐만 아니라 분석된 데이터들을 보면 월별 방문한 방문자 수, 사용한 브라우저, 가장 먼저 방문하는 페이지 등 소비자들의 출입과 움직임에만 관심을 가진다는 한계가 있다. 그리고 분석된 결과 확인을 위해서 생성되는 파일이나 웹을 참조해야하는 불편함이 있다.

이러한 문제점으로 인해 기존의 전자상거래 시스템들이 생성하고 분석하는 자료들은 소비자의 구매 패턴을 분석하기는 쉽지 않고 사용하기에도 불편하다.

그러므로 이를 분석하기 위한 로그를 생성하고 이러한 로그를 효과적으로 분석하며, 분석된 결과를 관리자가 확인하기 용이한 형태로 제시하는 새로운 시스템이 필요한 것이다.

## 2.2 OLE (Object-Lining and Embedding) DB

OLE DB는 ODBC를 아래의 두 가지 면에서 확장한 것이다[4].

- ▶ 데이터베이스 프로그래밍을 위해서 COM 인터페이스와 OLE를 제공
- ▶ 관계형과 비관계형 데이터 소스 양쪽을 위한 인터페이스를 제공

OLE DB의 구조는 그림 1과 같다.

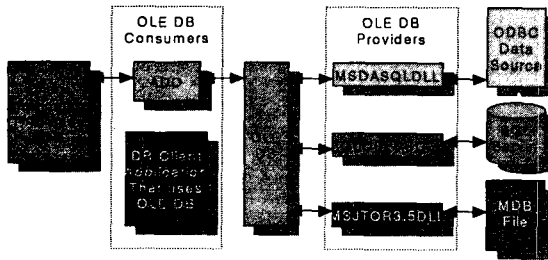


그림 1. OLE DB의 구조.

## 2.3 ADO (ActiveX Data Objects)

ADO는 OLE DB위에서 작동하게 되어있다. 즉 [그림 1]의 OLE DB consumer이다. 이러한 ADO를 사용하는 애플리케이션은 OLE DB 인터페이스를 직접적으로 사용한다. 그러나 ADO는 OLE DB와 비교 시 훨씬 더 간단하고, 프로그래밍 언어같이 사용된다. 그리고 데이터 모델은 DAO의 오브젝트 모델과 유사하다. 즉 데이터 모델은 객체 지향적인 프로그래밍이 가능한 모델이다.

이상에서 살펴보면 ADO는 객체 지향 프로그래밍을 위한 오브젝트 모델을 가장 잘 지원하고, 관계형·비관계형에 상관없이 모든 소스의 지원이 가능하며, 수

행능력이나 하위레벨 제어 등의 면에서 보았을 때 다른 기술들과 비슷한 성능을 보인다. 이러한 장점들 때문에 본 논문에서는 데이터베이스 접근 시 ADO 기술을 사용하였다.

## 2.4 DBMiner 2.0

DBMiner는 OLAP(On-Line Analysis Processing) 마이닝 시스템이고 사용자에게 친숙한 데이터 마이닝 환경을 제공한다. 이것은 데이터 웨어하우스에서 빠른 속도로 마이닝 할 수 있고 다양한 기능을 제공한다. 또한 적합성 분석과 추론, 전략적 관리, 소비자와의 관계 및 법인 재산의 관리, BPR(Business reengineering)과 전체 품질 관리 등에 사용하는 것이 적합하다.

DBMiner는 효율적이고 효과적으로 이미 존재하는 데이터 큐브로부터 몇 가지 종류의 지식을 찾을 수 있다. 그 이외의 장점을 들면 다음과 같은 것을 들 수 있다.

- ▶ 많은 데이터 큐브를 다룰 수 있다.
- ▶ dicing, slicing, pivoting, drilling등의 연산을 통합하여 사용한다.
- ▶ 보안의 문제를 통합하여 처리한다.
- ▶ 데이터나 정보에 대한 가시화에 그래픽을 사용한다.
- ▶ 강력한 데이터 마이닝 쿼리 언어를 사용한다.
- ▶ association, classification과 cluster 마이닝 등을 수행할 수 있다.
- ▶ 빠르고 효율적으로 수행하며 반응한다.

## 3. 전자상거래에서의 소비자 패턴 분석 도구

올바른 전자상거래 사이트의 운영과 마케팅을 위해 소비자의 구매 패턴을 분석하는 것은 확실히 중요성을 가진다[5]. 그림 2는 본 논문에서 구현한 소비자 구매 패턴 분석 시스템의 흐름도를 보여준다.

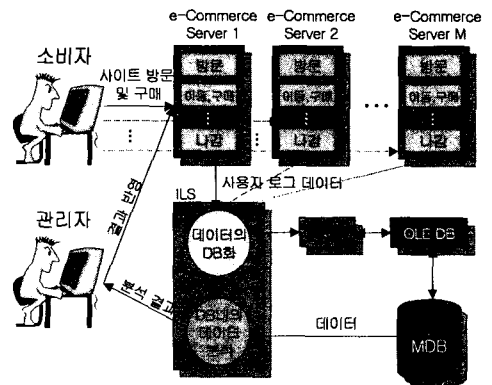


그림 2. 소비자 구매 패턴 분석 시스템의 흐름도

그림 2의 흐름을 설명하면 다음과 같다.

- step1 : 소비자가 사이트 방문 및 이동
  - step2 : e-Commerce Server가 step1의 방문 소비자 정보를 ILS(Intelligent Log Server:지능형 로그 서버) 시스템에 전송
  - step3 : ILS시스템은 e-Commerce Server로부터 받은 정보를 데이터베이스(MDB)에 저장
  - step4 : 소비자가 나갈
  - step6 : e-Commerce Server는 그 소비자의 세션이 종료되었음을 ILS에게 알림
  - step7 : ILS시스템은 전송받은 데이터를 저장
  - step8 : 저장된 정보를 DB Miner를 사용하여 분석
  - step9 : 분석된 결과를 반영하여 웹사이트를 운영
- 3.1절과 3.2절에서는 각 부분의 기능과 구동원리를 설명한다.

### 3.1 e-Commerce 서버

e-Commerce 서버는 전자상거래 사이트를 지원하는 서버로서 단 하나의 서버 시스템일 수도 있고 여러 개로 분산되어 운영되는 시스템일 수도 있다. 분산환경으로 운영되는 여러 서버를 각각 구별하고, 또 현재 소비자가 방문하고 있는 페이지가 어느 페이지인지, 그 페이지에서 얼마나 머무르고 있는지 등을 구별하기 위해서 각 페이지마다 ILS가 인식할 수 있는 태그(Tag)를 둔다.

소비자의 방문시 소비자의 IP, 방문 시간, 사용 브라우저 등의 정보를 우선 ILS에게 전송한다. 그 이후 각 페이지를 방문할 때마다 태그와 페이지 방문 시간을 ILS에 전송한다. 소비자가 방문을 마치고 사이트를 떠날 때 이 방문자에 의한 세션이 종료되었음을 알린다.

### 3.2 지능형 로그 서버(ILS)

지능형 로그 서버란 전자상거래 쇼핑물의 환경에서 사용자 로그를 데이터베이스화하여 저장하고 데이터베이스에 저장된 정보를 서버 종류별, 시간별, 페이지별 등으로 분석하는 서버를 말한다.

이 서버는 크게 아래의 2가지 모듈로 나눌 수 있다.

- 1) 데이터의 데이터베이스화 모듈
- 2) 데이터베이스내의 데이터 분석 모듈

1)의 모듈은 OLE DB provider 상에서 작동하는 ADO 기술을 사용하여 MDB를 생성하고, 데이터를 저장, 삭제, 업데이트 등이 가능하다. 표1은 OLE DB provider를 사용한 MDB를 생성하는 ADO 모듈의 일부이다.

표1 MDB 생성 ADO 모듈 소스

```
void CreateDatabaseX()
{
    HRESULT hr = S_OK;
    _CatalogPtr m_pCatalog = NULL;
    _bstr_t strConn("Provider=Microsoft.JET.OLEDB.4.0;");
```

2)의 모듈은 1)에서 저장된 데이터를 DBMiner2.0의 Associator 모듈이나 Classifier 모듈을 사용하여 분석하는 작업을 수행한다.

### 3.3 실험 및 평가

현 전자상거래 사이트들의 문제점으로 소비자 구매 패턴의 연구의 미비함을 들고 이를 해결하기 위한 해결책으로 지능형 로그 서버를 제안하였다. 이 서버는 정보의 데이터베이스화 시 ADO 기술을 사용함으로써 인체 데이터베이스 접근 시간을 단축하였다. 또 DB Miner2.0을 사용하여 소비자의 구매 패턴을 쉽게 분석할 수 있었다.

### 4. 결 론

본 논문에서는 기존의 전자상거래 시스템과 달리 로그를 데이터베이스에 저장하고 이를 분석하는 지능형 로그 서버를 두어 로그 정보를 분석하는 방법을 이용하였다. 이 시스템에서는 많은 시간이 소요되는 데이터베이스 접근 시간을 ADO를 사용하여 단축하였고, 웹을 통한 소비자 구매 패턴 분석 및 마케팅 전략 구축, 소비자 서비스 개선 등을 추구하였다.

앞으로는 지능형 로그 서버를 분산하여 운영할 수 있는 방법에 대해 연구하고자 한다.

### [참 고 문 헌]

- [1] HTTP://www.cnk.or.kr/front/
- [2] 이황규, 박규현, 이은령, "전자상거래의 개념에서 쇼핑물 추출까지", 2000
- [3] Jong Soo Park, "Data Mining : An Overview from Database Perspective", 1998
- [4] HTTP://www.microsoft.com/data/tch-vc21.htm
- [5] HTTP://db.cs.sfu.ca/WebMiner/
- [6] Osmar R. Zaiane, Man Xin, Jiawei Han, "Discovering Web Patterns and Trends by Applying OLAP and Data Mining Technology on Web Logs", 1998
- [7] Ming-Syan Chen, Jong Soo Park, "Efficient Data Mining for Path Traversal Patterns", 1998
- [8] Martine F. Arlitt, Carey L. Williamson, "Web Server Workload Characterization: The Search for Invariants (Extended Version)", 1996