

# RDBMS를 이용한 웹기반 다중항목 데이터 분석 시스템 설계 및 구현

□ 이선영\*, 이지선\*, 최정민\*, 장근\*, 이병수\*  
\*인천대학교 컴퓨터공학과  
e-mail : [sunlee77@incheon.ac.kr](mailto:sunlee77@incheon.ac.kr)

## A Design and Implementation of Web-based Multiple item Data Analysis System using RDBMS

Sun-Young Lee\*, Ji-Sun Lee\*, Jung-Min Choi\*, Keun Chang\*, Byoung-Soo Lee\*  
\*Dept. of Computer Science, Incheon University

### 요 약

기존의 오프라인에서 이루어지는 데이터 분석 방법은 대부분 수작업으로 이루어지기 때문에 데이터를 분석하는데 시간과 장소의 제약을 받으며 데이터 수집 과정에서 오류 발생률이 높게 나타나고, 예전보다 분석해야 할 데이터의 양이 방대해짐에 따라 사용자는 더 많은 시간과 노력을 할애해야 한다. 이에 데이터를 신속, 정확하게 분석하기 위해서는 보다 향상된 데이터 분석 기법의 개발이 필요하다. 따라서 본 논문에서는 이러한 문제점을 해결함과 동시에 온라인 환경에서 사용자가 데이터를 효율적으로 입력하고 분석할 수 있는 다중항목 데이터 분석 시스템을 제안한다. 이 시스템으로 사용자는 현재 일반화되어 있는 웹을 통하여 분석된 데이터 결과를 제공받을 수 있다.

### 1. 서론

웹은 고성능 컴퓨터의 보급과 데이터베이스 이용이 증가하면서 해마다 정보의 양이 급증하고 있으며, GUI 환경의 웹브라우저로 일반인도 편리하게 웹상의 정보를 제공받고 있다. 이렇게 제공되는 웹상의 방대한 정보는 대부분 데이터베이스에 저장되어 제공되어 지는데 최근 대내외적으로는 이러한 방대한 데이터의 입력과 데이터 분석 기법의 개발이 활발히 연구되고 있다. 이러한 이유는 기존의 오프라인에서 이루어지는 데이터 분석 방법이 대부분 직접 수작업으로 이루어지기 때문에 데이터를 분석하는데 많은 시간과 노력이 소요되며, 장소의 제한과 데이터 수집 과정에서 오류 발생률이 높게 나타나는 문제점이 있기 때문이다. 따라서 본 논문에서는 사용자가 수작업으로 일관된 데이터 수집과 분석의 문제점을 현재 일반화되어 있는 웹을 이용하여 시간과 장소에 구애받지 않고 데이터를 분석할 수 있는 'RDBMS를 이용한 웹기반 다중항목 데이터 분석 시스템'을 설계, 구현하였다. 이 시스템은 클라이언트에 있는 사용자가 웹 브라우저를 통해 인터넷에 접속하여 데이터를 입력하면 데이

터베이스 서버에 저장되고 분석엔진에서 데이터를 분석 처리한다. 분석된 결과는 데이터베이스 서버에 저장되며, 결과를 웹으로 출력하게 된다.

이와 같은 방법은 사용자의 요구에 따라 결과를 신속, 정확하게 확인할 수 있으며 그래프나 도표 유형으로 출력하여 쉽게 데이터를 분석할 수 있다. 또한 데이터베이스 서버에 결과가 누적되므로 필요한 데이터를 비교 분석할 수 있다.

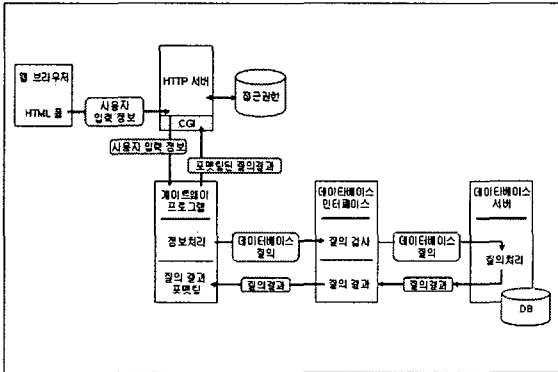
본 논문은 실제적인 구현을 목표로 다음과 같이 진행된다. 2 장에서는 웹과 데이터베이스의 연동에 관하여 기술한다. 3 장에서 본 논문에서 구현한 분석엔진을 설계하고, 4 장에서는 실제 구현한 사례를 제시한다. 마지막으로 5 장에서 결론을 맺고 향후 연구과제에 대해 기술한다.

### 2. 웹과 데이터베이스의 연동

웹 서버와 데이터베이스 서버의 연동은 분석해야 할 주요 데이터들이 대부분 데이터베이스에 저장되어 있기 때문에 두 시스템의 연동 기술은 매우 중요하다.

따라서, 웹 서버와 외부 프로그램을 연동하기 위한 인터페이스로서 CGI(Common Gateway Interface)를 이용하여 웹과 데이터베이스를 연동할 수 있다. 즉, 웹 서버가 전송한 질의 정보를 데이터베이스 엔진으로 전달하고 HTML 형태의 실행 결과를 반환받게 된다. 이렇게 반환받은 결과를 클라이언트로 출력할 수 있는 CGI 프로그램을 작성하면 웹과 데이터베이스의 연동이 가능하다.

일반적으로 웹브라우저와 데이터베이스 서버의 연동 구조는 [그림 1]과 같다.



[그림 1] 웹과 데이터베이스의 연동 구조

또한 웹과 데이터베이스를 연동하는 메커니즘 중에는 최근 웹 응용에 많이 사용되는 자바서블릿, ASP, PHP를 이용한 데이터베이스 연동이 있다.

### 2.1 자바서블릿을 이용한 데이터베이스 연동

서블릿(Servlet)은 서버쪽에서 실행되는 애플릿으로, 자바서버 API(Application Programming Interface)의 서브셋이다. 서블릿은 CGI와 유사하게 HTML 폼이나 클라이언트 애플릿을 통해 PUT/POST 메소드로 사용자 질의를 전달받고, JDBC를 통해 데이터베이스 서버에 전달하여 실행 결과를 반환받게 된다. 이렇게 반환받은 결과를 HTML 문서로 포맷팅한 다음 클라이언트로 전달하면 데이터베이스에 연결된 것이다.

자바로 작성된 프로그램이기 때문에 어느 플랫폼에서나 수행될 수 있으며, 용이한 데이터베이스 접근과 자바애플릿과의 손쉬운 통합 기능을 제공한다.

### 2.2 ASP를 이용한 데이터베이스 연동

ASP(Active Server Page)는 동적인 대화형 웹페이지를 구성하며, 웹 서버 응용 프로그램을 작성하고 실행하기 위해 사용할 수 있는 서버측 스크립트 언어이다. OLE DB를 기반으로 한 ADO(Active Data Object) 컴포넌트를 이용하여 SQL 문장을 전송하고, 그 결과 레코드셋을 받아서 HTML로 변환하는 루틴이 서버측에서 수행되면 데이터베이스에 연결된 것이다.

ASP는 배우기가 쉽고 데이터베이스 연동도 편리해서 많이 사용되지만 윈도우 운영체제에서만 동작하는

한계점이 있다.

### 2.3 PHP를 이용한 데이터베이스 연동

PHP(Professional HTML Preprocessor)는 사용자가 HTML 폼을 통해 입력한 값을 웹 서버 상에서 처리한 후 그 결과를 HTML과 같은 형태로 가공하여 사용자의 브라우저에 전달하는 서버측 스크립트 언어이다. 사용자 질의가 웹 브라우저를 통하여 웹 서버로 전달되면, 서버측의 PHP가 그 질의를 해당 데이터베이스 서버에 대한 확장 모듈을 통하여 데이터베이스 엔진으로 전송하게 된다.

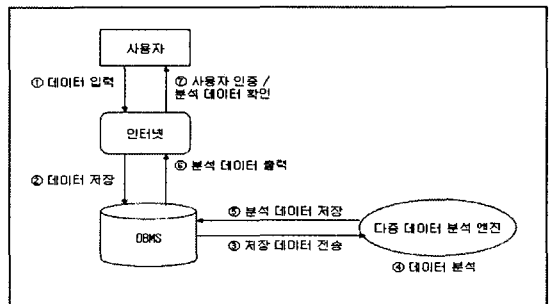
PHP는 속도나 개발 편의성, 여러 확장 기능이 있어서 ASP보다 사용하기가 어렵고 복잡하다. 그러나 윈도우 운영체제뿐만 아니라 리눅스나 유닉스 등 매우 광범위한 운영체제에서 동작하고, 대부분의 데이터베이스와 공개된 각종 프로그램을 지원한다. 또한 별도의 프로세스를 생성하지 않기 때문에 실행 속도가 빠르며, 최근 배포된 PHP4는 Zend라는 해석엔진을 채용하여 수행 속도를 향상하였다.

본 논문에서는 이러한 특징을 고려해서 PHP4 Zend를 이용하여 데이터베이스와 연동하였다.

### 3. 분석엔진 설계

본 논문에서는 웹 상에서 다중항목의 데이터를 분석하는 시스템을 구현하는데 목적을 두고 분석엔진을 설계하였다.

[그림 2]는 전체적인 구성도를 나타내고 있다. 이 시스템은 사용자가 웹 상에서 분석하고자 하는 데이터를 입력하면 입력된 데이터를 데이터베이스에 저장하고, 분석엔진에서 데이터를 항목별로 처리하여 결과를 웹을 통해 제공한다.



[그림 2] 전체적인 구성도

이 시스템은 아래와 같은 단계로 수행된다.

- ① 데이터 입력 : 사용자가 인터넷에 접속하여 웹 상에서 분석하고자 하는 데이터를 입력한다.
- ② 데이터 저장 : 입력된 데이터는 데이터베이스 서버에 저장된다.
- ③ 저장 데이터 전송 : 저장된 데이터를 분석엔진으로 전송한다.

- ④ 데이터 분석 : 다중 데이터 분석엔진에서 데이터를 항목별로 분석 처리한다.
- ⑤ 분석 데이터 저장 : 분석된 데이터는 데이터베이스 서버에 재저장한다.
- ⑥ 분석 데이터 출력 : 이렇게 분석된 데이터를 웹으로 출력한다.
- ⑦ 사용자 인증 / 분석 데이터 확인 : 사용자는 인증 절차를 거친 후 분석된 데이터를 확인한다.

여기서 분석엔진의 구현을 위하여 시스템에 Linux Kernel 2.2.17 과 Apache 1.3.20 를 탑재하고, 대용량의 데이터를 저장 및 관리하기 위하여 Oracle 8.1.5 DBMS 를 사용하였다. 웹 서버 구축에 따른 사용자 인터페이스로 PHP 4 Zend 를 사용하여 구현하였으며, 데이터 분석은 Visual Basic 6 과 SPSS 10.0 을 사용하여 처리하였다.

#### 4. 구현 사례

##### 4.1 원격 모의고사 관리 시스템

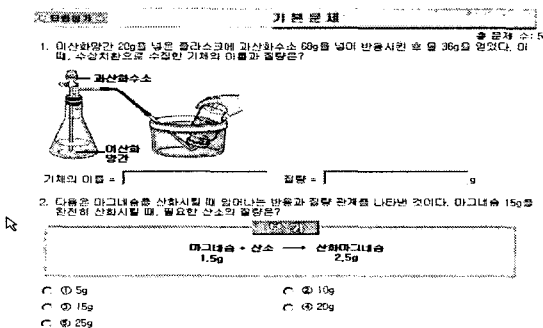
다중항목 데이터 분석 시스템은 다양한 용도로 응용될 수 있다.

본 논문에서는 그 중 한가지 예로 현재 오프라인에서 이루어지고 있는 시험 진행에서 성적 분석 처리까지의 전과정을 온라인 상에서 처리할 수 있는 원격 모의고사 관리 시스템을 구현하였다. 이 시스템은 온라인상에서 시험을 보고 채점된 결과를 데이터베이스 서버에 저장한다. 관리자는 네트워크에서 통계 분석 엔진으로 데이터를 분석 처리한 후 결과를 데이터베이스 서버에 재저장한다. 인증 절차를 거친 사용자는 웹을 통해 결과를 제공받는다.

##### 4.2 사용자 인터페이스 구현

사용자 인터페이스는 웹을 통해 결과를 제공받을 수 있는 사용자 인터페이스와 다중항목 데이터 분석을 수행할 관리자 인터페이스로 나누어진다.

먼저 사용자 인터페이스는 PHP4 를 사용하여 구현하였으며, 인증 절차를 거쳐 웹상에서 문제를 풀고 정답을 제공받는다. [그림 3] 은 사용자의 정답 데이터를 입력받기 위한 인터페이스를 보여준다.



[그림 3] 사용자 인터페이스

웹을 이용한 사용자 인터페이스는 별도의 실행 프로그램없이 인터넷에서 결과를 확인할 수 있다.

반면, 클라이언트를 이용한 관리자 인터페이스는 데이터베이스 서버에 저장되어 있는 데이터를 가져오고 다양한 항목의 데이터 분석 결과를 산출한다. 산출된 결과는 데이터베이스 서버에 재저장함으로써 인증 절차를 거친 후 인터넷에서 결과를 확인할 수 있다.

##### 4.3 데이터 분석 시스템 구현

관리자에 의해 이용되는 데이터 분석 시스템은 Visual Basic 과 SPSS 를 연동하여 구현하였으며, 데이터베이스에 접속하여 데이터를 가져오고 처리된 분석 결과를 재저장하는 기능을 가진다.

[그림 4]는 ADODB 객체를 이용하여 데이터베이스와 연동하는 것을 보여준다.

```

'선언부
Dim Db As ADODB.Connection
Dim Rs As ADODB.Recordset

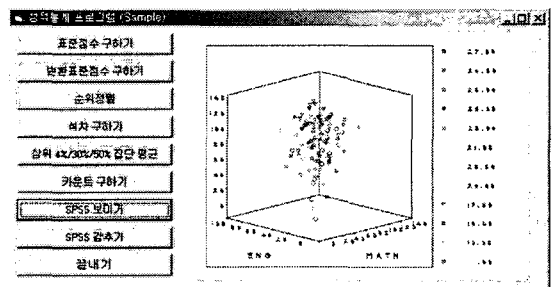
'DB 열기
Db.Open "Provider=MSDAORA.1;Password=UserPassword;Persist
Security Info=True;User ID=UserID;Data
Source=ScoreData"

'Query 데이터 레코드셋으로 불러오기
krs = "select * from score_board"
Rs.Open krs, Db

'끝낼때 반드시 닫는다.
Rs.Close
Db.Close
Set Rs = Nothing
Set Db = Nothing
    
```

[그림 4] 데이터베이스 연동

[그림 5]는 다양한 항목의 데이터 분석 처리 과정에서 과목별 성적 분포도 결과를 보여준다.



[그림 5] 데이터 분석 처리 과정

Visual Basic 에서의 SPSS 연동은 OLE 객체를 사용하며, 생성된 객체에 각각의 데이터 분석 코드를 삽입하면 된다.

```

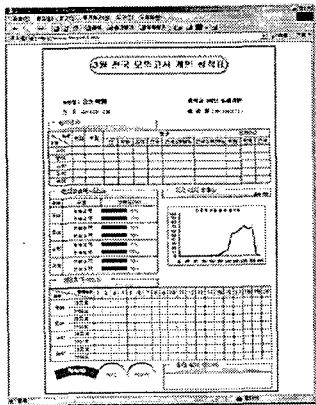
Screen.MousePointer = vbHourglass

' SPSS 연동
Set objSpssApp = CreateObject("SPSS.Application")
Set objDataDoc = objSpssApp.OpenDataDoc(App.Path + "\data.sav")
Set objOutPutDoc = objSpssApp.NewOutputDoc
    spss_run = True
    While objSpssApp.IsBusy
        Screen.MousePointer = vbHourglass
    Wend

' 입력 데이터를 분석하여 석차 산출
strRANK = "rank VARIABLES=sum (D)/RANK /PRINT=YES /TIES=LOW ."
strRANK1 = "rank VARIABLES=kor (D)/RANK /PRINT=YES /TIES=LOW ."
strRANK2 = "rank VARIABLES=eng (D)/RANK /PRINT=YES /TIES=LOW ."
strRANK3 = "rank VARIABLES=math (D)/RANK /PRINT=YES /TIES=LOW ."
strRANK4 = "rank VARIABLES=phy (D)/RANK /PRINT=YES /TIES=LOW ."
strCommand = " SAVE TRANSLATE OUTFILE=" + App.Path + "\db.dbf"
strCommand = strCommand + " /TYPE=DB3 /MAP/REPLACE."
    
```

[그림 6] 석차 산출 모듈

이렇게 분석된 데이터는 백업을 위해 클라이언트 서버에 저장되는 동시에 데이터베이스 서버로 전송된다. 다중 데이터 분석 시스템을 거친 데이터는 인터넷을 통한 인증 과정을 거쳐 [그림 7]과 같은 형태로 제공되어진다.



[그림 7] 구현 결과

### 5. 결론 및 향후 연구과제

본 논문에서는 대부분 수작업으로 이루어지는 데이터 분석 방법을 웹상에서 데이터를 입력하고 입력한 데이터를 분석하는 'RDBMS'를 이용한 웹기반 다중항목 데이터 분석 시스템'을 설계, 구현하였다. 본 시스템은 기존의 오프라인 방법과는 달리 시간과 장소에 구애 받지 않고 효율적으로 데이터를 분석할 수 있는 시스템이다.

본 시스템이 다중항목 데이터를 효과적으로 분석할 수 있는지를 검증하기 위하여 원격 모의고사 관리 시스템에 적용하여 구현해 보았는데, 그 결과 상당히 만족스러운 결과를 얻을 수 있었다.

본 시스템은 웹을 기반으로 개발된 시스템이기 때문에, 인터넷 접속이 가능한 어느 곳에서나 웹 브라우저를 이용하여 편리하고 쉽게 데이터를 분석할 수 있다. 분석 결과는 인터넷에서 도표나 그래프로 제공되므로 비교 분석이 가능하여 변화치를 뚜렷하게 파악할 수 있는 장점을 가지고 있다.

그러나, 네트워크를 이용한 데이터 전송에서 트래픽이 발생되므로 이 문제를 근본적으로 해결할 방안이 필요하다.

향후 연구과제로는 분산 시스템을 이용한 미들웨어를 도입하여 보다 안정적이고 효율적인 시스템을 구축하는 것이며, 또한 본 시스템을 기반으로 데이터 분석을 위한 다양한 기법을 다른 시스템에 적용하면 여러 분야에서 유용하게 활용될 것으로 기대된다.

### 참고문헌

- [1] AI Dev, " Database-Sql-Rdbms Howto:Postgresql Object Relational Database System"
- [2] Henderson, Rebecca Sue.Technology, " The design and use of relational databases", University of Washington, 1999.
- [3] 김성식, " 오라클 연동을 위한 PHP 웹프로그래밍", 대청미디어 pp.11-14, 1999.
- [4] 구홍서, " WWW 과 데이터베이스 연동기술의 조사 분석", 한국정보과학회지 제 18 권 제 4 호, pp.32-40, 2000.
- [5] 서재일 외 2, " Web 상에서 객체 지향 설계를 이용한 전자문서 수발 시스템 개발", 한국정보처리학회 추계학술발표논문집 제 5 권 제 2 호, pp.37-42, 1998.
- [6] 이경자 외 1, " WWW 상에서 JAVA 와 데이터베이스를 이용한 통계정보시스템의 구현", 한국정보처리학회 추계학술발표논문집 제 5 권 제 2 호, pp.163-166, 1998.
- [7] 이승호 외 3, " 웹 브라우저와 음성 인터페이스 연동기법에 관한 연구", 한국정보처리학회 추계학술발표논문집 제 5 권 제 2 호, pp.209-212, 1998.
- [8] 이지선 외 3, " 웹 기반 원격 모의고사 관리 시스템 설계 및 구현에 관한 연구", 한국정보처리학회 추계학술발표논문집 제 7 권 제 2 호, pp.493-496, 2000.
- [9] 차상균 외 2, " WWW 상에서 지리정보 데이터베이스를 위한 사용자 인터페이스 설계 및 구현", 한국정보과학회 춘계학술발표논문집 Vol.23, No.1, pp.143-146, 1996.