한국항행학회 논문지 제12권 제1호 2008년 2월

상용 도구를 이용한 대용량 교통 데이터의 추출 및 적재 방안

A Method for Extraction and Loading of Massive Traffic Data using Commercial Tools

우찬일*, 전세길**

Chan-Il Woo*, Se-Gil Jeon**

요 약

지능형 교통 체계(ITS: Intelligent Transport Systems)는 기존의 교통에 정보통신, 전기 및 전자, 기계, 제어 등의 기술을 접목시켜 도로 및 교통 체계의 안전 및 운영 효율을 극대화하여 교통 문제를 해결하는 시스템이다. 현재 대량의 교통 원시 데이터를 위해 시스템을 통합하는 문제는 다른 소스데이터 간의 표현상의 이질성 및 데이터의 양과 다양성 그리고 소스의 자동화와 서로 다른 성능 등 몇 가지 문제에 직면해 있다. 본 논문에서는 고속도로 교통관리시스템과 우회도로 교통정보시스템에서 수집되는 이질적인 대용량 교통데이터의 데이터 웨어하우스를 구축하기 위해 상용 도구를 활용해 추출 및 적재하는 방안을 기술한다. 또한, 고품질 교통데이터를 제공하기 위한 통합 품질관리 기법을 적용하여 평가 한다.

Abstract

The ITS(Intelligent Transport System) enables us to provide solutions on traffic problems, while maximizing safety and efficiency of road and transportation systems, by combining technologies from information and communication, electrical engineering, electronics, mechanics, control and instrumentation with transportation systems. The issues that an integration system for massive traffic data sources must face are due to several factors such as the variety and amount of data available, the representational heterogeneity of the data in the different sources, and the autonomy and differing capabilities of the sources. In this paper, we describe how to extract and load of the heterogeneous massive traffic data from the operational databases, such as FTMS and ARTIS using commercial tools. Also, we experiment on traffic data warehouses with integrated quality management techniques for providing high quality data.

Key words : ITS, Massive traffic data, ETL, OWB

I. 서 론

현재 고속도로나 우회도로 등 교통 분야에서 취득되는 매일 수 GB에 달하는 교통 원시데이터는 활용가치가 매우 높은 데이터들이다. 이러한 교통 원시데이터의 종류로는 차량검지기(VDS : Vehicle Detection System) 데이터, 통행료 수납시스템 데이터(TCS : Toll Collection System), 하이패스(Hi-Pass) 데이터 및 기타 교통 운영에 관련되는 데이터가 있으며, 대부분의 용량을 차지하는 것은 차량검지기 데이터이다[1].

* 서일대학 정보통신과(Dept. of Information and Communication Eng., Seoil College)
** 한국도로공사 도로교통기술원(Korea Highway Corporation HTT(Highway and Transportation Technology Institute))
· 제1저자 (First Author) : 우찬일
· 접수일자 : 2008년 1월 9일

이와 같은 교통 원시데이터 들을 수집하여 연구에 활용할 경우, 고속도로 교통소통 대책 수립, 안전성 향상, 명절 및 특송 기간 교통수요분석, 돌발 상황에 따른 교통정체예측, 데이터처리 모형 및 알고리즘 개선 등의 분석에 활용할 수 있다. 또한, 이러한 교통 이력데이터를 저장하고 관리하기위한 연구 분야인 교통 데이터웨어하우스에 대한 연구는 국제학술계에서 하나의 분야로서 인정될 만큼 최근 그 중요성을 더해가고 있다. 이와 같은 분야를 선도하고 있는 미국의 경우, 2000년대부터 교통이력 데이터 관리시스템(ADMS : Archived Database Management System) 관련 기술과 시스템 개발에 노력하고 있다[2]-[4]. 우리나라의 경우, 교통정보 수집 인프라 구축 및 데이터 처리 프로세스와 관련된 모형이나 알고리즘 개발에 치중되어 교통데이터 웨어하우스 구축이나 활용에 대한 노력이나 연구는 진행된 바가 없다. 또한, 이와 같은 데이터는 실시간 교통정보 제공에 국한되어 이용되고 있을 뿐 교통 분석 등의 용도로는 활용되고 있지 않은 실정이다. 이러한 대용량 교통데이터 웨어하우스에서의 최대 관건은 대용량 교통데이터 저장을 위한 물리적, 논리적 기반 환경을 구축하고 관련된 시스템으로부터 교통데이터를 어떻게 통합하느냐에 달려 있다고 볼 수 있다. 그리고 통합된 교통데이터 웨어하우스는 통합이 수행되면서 여러 가지 원인으로 인하여 많은 오류를 가지고 있어 어떻게 오류를 해결하고 분석이 용이한 형태의 고품질 데이터를 제공하느냐가 가장 중요한 사항이다.

본 논문에서는 고속도로 교통관리시스템과 우회도로 교통정보 시스템을 대상으로 교통데이터 웨어하우스를 구축하기 위한 데이터 추출과 적재 방안을 유형별로 기술하고 오라클 OWB(Oracle Warehouse Builder)를 활용하여 추출과 적재를 수행한다[5]. 또한, 고품질의 교통데이터를 제공하기 위해 기존에 제안된 통합 품질관리 프레임워크를 적용하여 운영계 시스템으로부터 효율적으로 데이터를 통합하기위한 방안을 기술하고 OWB를 이용하여 추출, 적재한 경우와 통합 프레임워크를 이용하여 적재한 경우를 비교 평가한다.

Ⅱ. 관련 연구

2-1 ADMS

ADMS는 교통관리시스템에서 수집되는 데이터를 다양한 목적으로 효율적인 이용이 가능하도록 데이터의 가공처리 및 저장 그리고 공급과정을 체계적으로 구성하여 1, 2차 데이터처리과정(데이터 처리 및 검색)과 데이터 저장기능을 갖는 물리적 시스템으로 미국 National ITS Architecture에서는 ADMS를 ITS의 서브시스템으로 정의하고 있다. ADMS 관련 연구는 미국을 중심으로 2000년 이후 활발하게 진행되었다. 그러나 아직까지 데이터 웨어하우스를 통합 구축하여 다양한 분석환경을 제공하기 위한 방안이 종합적으로 연구되고 있지 않아 대용량 교통 이력데이터에 대한 설계 방법론과 구축 방법을 제시하지 못하고 있으며, 교통 이력데이터 시스템을 효율적으로 구성하기위해 구체적인 실 사례를 이용해 물리적, 논리적 설계 방안을 제시한 사례가 부족한 상태이다.

2-2 Oracle Warehouse Builder

OWB가 수행하는 주요 기능은 데이터 마이그레이션 분리된 데이터 소스로부터의 데이터 품질 정보 제공을 위한 정제와 연관된 메타데이터의 설계 및 관리이다.

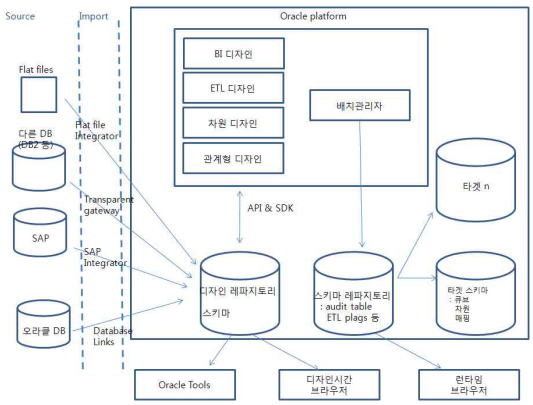

그림 1. OWB 아키텍쳐
Fig. 1. OWB architecture.

오라클 OWB의 장점으로는 오라클 10g와의 호환성이 좋고 상대적으로 다른 툴에 비하여 사용 방법이

용이하며, 가격이 저렴하다. 반면에 빈약한 레퍼런스 때문에 개발 시에 어려움이 있을 수 있다.

2-3 교통데이터 웨어하우스

본 논문에서는 고속도로 교통관리 시스템(FTMS : Freeway Transportation Management System)과 우회도로 교통정보 시스템(ARTIS : Alternative Road Transportation Information System)에서 제공되는 다양한 교통데이터를 대상으로 교통데이터 웨어하우스를 구축하고자 한다. 교통데이터 웨어하우스는 분산된 데이터 소스들로 부터 정보를 추출, 변환, 적재(ETL : Extract, Transform and Load)한 후 사용자들에게 일관된 정보를 제공하기 위해 중앙 집중적인 저장소에 통합 저장하여 데이터에 대한 다양하고 복잡한 분석을 가능하게 한다. 따라서, 교통데이터 웨어하우스에서는 데이터의 통합을 위한 추출과 변환 방법에 대한 연구가 필요 하며, 데이터를 저장하기 위한 데이터 모델은 확장이 용이하고 다양하고 뛰어난 검색 기능을 지원하도록 설계해야 한다. 교통데이터 웨어하우스 관리 시스템은 그림 2와 같이 FTMS와 ARTIS로부터 모든 데이터를 수집하여 중앙 집중적으로 저장하는 형태로 저장된 데이터는 다양한 형태로 가공, 평가, 분석될 수 있다.

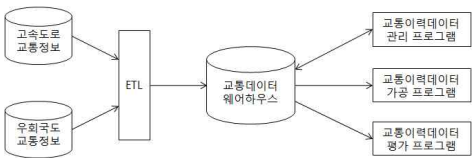

그림 2. 교통 데이터 웨어하우스 관리 시스템
Fig. 2. Traffic data warehouses management system.

실시간 정보를 제공하는 FTMS나 ARTIS는 운영 시스템의 부하가 크므로 교통데이터 웨어하우스는 운영 시스템의 활용이 적은 일정 시간마다 데이터를 추출하는 방식을 취하며. 데이터를 추출한 후 데이터의 품질상태에 따라 추가적으로 정제하는 과정이 필요할 수 있다[6]-[8]. 현재 다루고자 하는 교통 관련 데이터는 여러 가지 가공 프로세스를 통해 오류 보정이나 집계 값 계산을 수행한 후 다양한 연구에 활용이 되고 있는데 이러한 가공 프로세스를 교통데이터 웨어하우스의 정제 과정에 적용한다. 가공된 데이터는 평가 프로그램을 통해 완전성, 유효성 등이 평가되며, 교통데이터 웨어하우스는 교통이력데이터 관리 프로그램을 통해 사용자들에게 제공된다[9],[10].

Ⅲ. 교통데이터 추출, 적재 및 평가

3-1 교통데이터 추출

관련된 소스로부터 데이터를 추출하기 위해서는 데이터를 추출하는 시스템의 현재 시스템 자원 여유율과 여유 공간을 고려하여 방법을 선택해야 한다. 일반적으로 데이터 웨어하우스에서 많이 사용하는 방법으로는 ETCL(Extraction Transformation Cleansing Loading) 도구를 활용하는 방법이 있으며, 표 1에서와 같이 Case 1과 Case 2가 이 경우에 해당된다.

표 1. 운영계 데이터 추출 방법
Table 1. Extraction methods from operational data.

구분	추출 방법
Case 1	ETCL 도구를 활용해 DB에서 직접 추출
Case 2	ETCL 도구를 활용해 운영계에서 SAM 형태로 추출
Case 3	운영계에서 export 도구를 사용해 추출
Case 4	OS 복사로 DB 데이터 파일 추출

Case 1은 DB에서 직접 칼럼별로 추출하는 방법으로 사용하기 쉬운 반면에 약 1백만건 미만인 경우에 사용할 수 있으며 소스 시스템에 많은 부하가 걸리는 단점이 있다. Case 2는 SAM(Sequential Access Method) 파일을 만들어 추출하는 방법으로 소스 시스템에 부하를 상당히 줄일 수 있는 장점이 있지만 작업의 단계가 많아지고 소스 시스템에 충분한 공간이 있어야 한다. Case 3과 Case 4의 경우도 Case 1에 비해서 상당부분 소스 시스템의 부하를 줄일 수 있지만, 소스 시스템에 충분한 공간이 있어야 하며 만들어진 데이터는 동일한 DBMS에서만 적재될 수 있다는 단점과 데이터를 전송 할 수 있는 방법에 대한 네트워크 상황을 고려해야 한다. OWB를 활용하지 않은 경우의 고속도로 교통 데이터는 소스데이터의 특

징과 서버부하 문제를 고려하여 Case 3의 방법을 사용하고, 우회국도 데이터는 Case 4를 활용한다.

본 논문에서의 Case 1과 Case 2에서의 OWB를 이용한 데이터 추출은 지연 데이터 추출 기법을 사용하였으며 그 중에서도 날짜와 타임스탬프에 근거한 획득 방법을 적용하였다. 이 방법은 원시 시스템들의 정규 작업 외에 데이터 추출이 수행되어 운영계 시스템에 최소한의 영향을 미치는 장점을 가지고 있다. 그러나 현재 추출 시점에서의 소스 데이터의 상태로부터 이전 추출 시점에서의 상태와 비교하여 변경을 추적하므로 중간의 변화를 획득할 수 없는 단점이 있다.

3-2 교통데이터 적재

원시 데이터베이스로 부터 전송된 데이터를 목표 시스템의 데이터베이스에 적재하기 위해서는 표 2와 같은 방법을 고려할 수 있다.

표 2. 데이터 웨어하우스의 적재 방법
Table 2. Loading methods to traffic data warehouse.

구분	적재 방법
Case 1	ETCL 도구를 활용해 DB to DB로 적재
Case 2	SAM 파일을 ETCL 도구로 변환 후 DBMS의 AUTOLOADER를 이용하여 적재
Case 3	운영계에서 만들어진 DB Dump를 import 도구를 사용하여 적재
Case 4	DB 데이터 파일을 데이터베이스 복구 방법을 이용하여 적재

Case 1은 DB에서 DB로 직접 적재하는 방법으로 타겟 시스템에 부하가 상당히 걸리는 방법이다. Case 2는 SAM 파일을 이용하여 적재하는 방법으로, SAM 파일을 데이터베이스에 적재하기 위해서는 별도로 DBMS에서 제공하는 로더를 사용해 한다. 고속도로 데이터의 경우 초기에 적재된 백업 데이터는 확장자가 dmp인 오라클 덤프파일로 만들어져 있어 FTP로 교통 이력데이터 시스템으로 전송한 후 오라클 import 도구를 사용하여 적재한다. 변경되는 데이터의 기초정보는 초기에 적재된 정보의 변화가 있을 경우에만 전송하여 적재하고, 운영정보는 백업 주기를 고려하여 전송 후 적재한다. 우회국도 데이터는 운영체제에서 복사하는 방식으로 백업된 초기 및 변경 DB 파일을 전송하여 교통 이력데이터 시스템에 적재할 수 있다. 교통 데이터의 경우 데이터를 취득하는 기계 자체의 문제로 인한 오류나 데이터 전송 시의 문제로 근본적으로 많은 오류를 가지고 있다. 이러한 오류를 처리하는 대표적인 방법으로는 공간적으로 같은 위치의 이전 시간의 데이터를 활용해 결측을 처리하거나 주변 장치의 데이터를 이용해 보정하는 방법이 있는데, 본 논문에서 제안한 방법에서는 이러한 교통 관련 데이터 가공 프로세스를 데이터베이스 추출, 적재 과정에 프로시져 같은 형태로 작성하여 함께 수행하도록 하였다.

```
/*통합 데이터처리 프로세스*/
운영계 데이터베이스로부터 대상 원시데이터 추출 및 적재 수행

IF 검색 조건 입력 = True THEN
   대상 원시데이터의 위치 확인
   사용자의 권한 검사
   IF 사용자 권한 유효한가 = True THEN
     IF 시공간 검색 = True THEN
       시공간 검색 조건에 따라 대상 데이터 검색
       WHILE 대상데이터 != EOF
       교통공학 관점의 오류판단 수행
       교통공학 관점의 오류보정 수행
       IF 교통공학 관점의 오류판단& 오류보정 문제점 없음 = True
       THEN
         DB 관점의 데이터 값 오류 판단 수행
         DB 관점의 데이터 값 오류 보정 수행
         IF DB 관점의 오류판단& 오류보정 문제점 없음 = True
         THEN
           평활화 수행
           시공간 집계 수행
           IF Permanent 저장 = True && 사용자 권한 적합 = True
           THEN
              Permanent한 DB 공간에 결과 데이터 저장
           ELSE IF Temporal 저장 = True THEN
              Temporal한 DB 공간에 결과 데이터 저장
           ELSE IF 파일 저장 = Ture THEN
              선택한 형식으로 파일에 결과데이터 저장
           END IF;
         END IF;
       END IF;
       END WHILE;
    END IF;
    ELSE IF 기타 조건검색 = True THEN
     시공간 검색 결과에 대한 데이터처리 프로세스와 동일작업 수행
    END IF;
   ELSE
     대상 데이터 접근불가 통지
   END IF;
END IF;
```

그림 3. 통합 데이터처리 프로세스
Fig. 3. Integrated data processing process.

3-3 교통 데이터 평가

교통 데이터평가는 데이터 값 측면에서 운영계 원시데이터 값에 대한 평가와 교통데이터 웨어하우스에 대한 분석데이터 평가로 구분할 수 있다.

그림 4. 운영 및 분석 데이터 값의 데이터평가
Fig. 4. Evaluation of traffic data.

본 논문에서는 그림 5와 같은 통합 데이터평가 프로세스를 제안하고 단순히 OWB 처리프로세스를 적용해 추출, 적재한 경우와 비교 평가한다. 데이터 평가는 데이터 처리를 수행하지 않은 차로별 원시 데이터에 대한 값을 기준으로, 데이터베이스 측면 6가지 항목과 교통공학 측면의 5가지 항목으로 평가를 수행한다. 또한, 데이터 구조에 대한 평가는 데이터베이스 측면에서만 엔티티, 관계, 속성 평가가 수행된다[9].

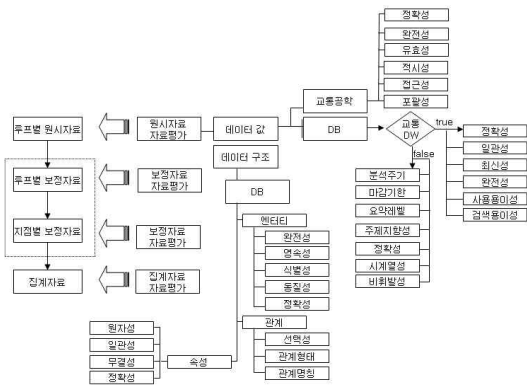

그림 5. 통합 데이터평가 프로세스
Fig. 5. Integrated process for traffic data evaluation.

데이터 처리를 포함한 데이터 추출 및 적재 후 데이터를 평가하기 위한 기준은 표 3과 같다.

표 3. 품질 평가 기준
Table 3. Criterion of quality evaluation.

품질	오류 데이터 항목
데이터 값	설계하고 다른 이상치 데이터 존재
	용어 정의가 일치하지 않음 규정, 표준이 일치하지 않음 속성 정의가 일치하지 않음 데이터 형식이 일치하지 않음
	누락된 데이터가 있음
엔티티 데이터 구조	두개미만의 속성을 가짐 두개미만의 인스턴스를 가짐
	키 역할을 하는 속성이 없음
	같은 엔티티에 대한 서로 다른 이름 존재 서로 다른 엔티티에 대한 같은 이름 존재 누락된 애트리뷰트 존재 테이블레벨의 제약조건 오류 존재

통합 프레임워크에서 데이터베이스의 품질 정확도(Qi)는 표 3을 참조하여 데이터 값이나 엔티티 데이터구조 측면의 추출 및 적재 전, 후의 오류 개수를 식 1에 대입하여 산출한다.

$$Q_i(\%) = w_j \sum_{j=1}^{m} S_{ij}, \qquad w_j = \frac{1}{m_i} \tag{1}$$

$$S_{ij} = \frac{1}{|a_{ij}|} \sum_{k=1}^{|a_{ij}|} (1 - \frac{n_{ijk}}{N})$$

S_{ij}: Q_i에 속하는 품질 부특성

$m_i = |S_{ij}|$: Q_i의 총 품질 부특성 개수

a_{ijk}: S_{ij}에 속하는 오류 데이터 항목 k

$|a_{ij}| = S_{ij}$에 속하는 총 오류 데이터 항목의 개수

N: 총 품질측정 대상 데이터 개수

n_{ijk}: a_{ijk}의 오류 데이터 개수

Ⅳ. 실험 및 결과

본 논문에서는 적재된 데이터의 품질 유효성을 보기 위해 기존에 통합품질 프레임워크와 OWB를 사용한 경우를 비교해 보았다. 실험 대상 데이터는 고속도로 경부선에 대한 검지기 데이터로 2007년 7월 1일부터 2007년 9월 30일까지 2개월간의 데이터에 대해 처리한다. 실험은 OWB의 디자인 센터에서 소스와

타겟 데이터베이스를 설정한 후, 소스와 타겟 내의 테이블간의 매핑은 컬럼 별로 설정해야 하며 이를 수행한 화면은 그림 6과 같다.

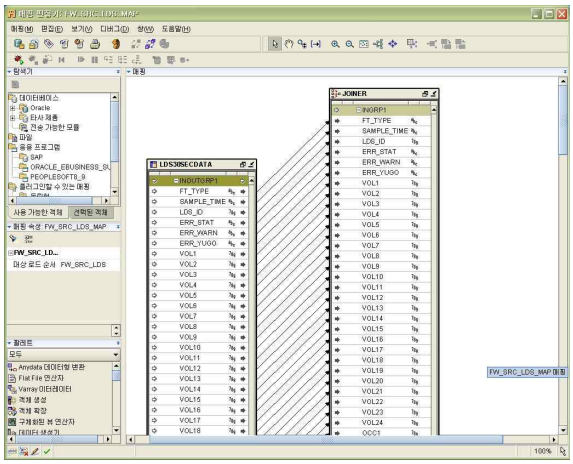

그림 6. 테이블 매핑 화면
Fig. 6. Screen of table mapping.

그림 7은 데이터 값에 대한 통합프레임워크와 OWB를 수행했을 때를 비교하기 위해 1개월과 2개월치 데이터에 대한 통합 시 오류가 발생해도 아무런 처리도 하지 않은 원시데이터와 OWB를 이용한 경우 그리고 본 논문에서 적용한 통합 프레임워크를 적용한 경우와 적용하지 않은 경우를 비교 평가한 것을 나타낸다. 그림 7에서 원시데이터는 앞서 기술한 데이터베이스 추출 및 적재 방안 중에서 Oracle OWB 처리도구는 단순히 case1에 데이터 처리 프로세스를 적용하여 추출 적재한 경우이며, 통합 프레임워크는 본 논문에서 제안한 통합프레임워크를 이용해 데이터 처리를 포함한 추출 적재 후 정확도를 평가한 결과이다. 비교 평가 결과 통합 프레임 적용 시 높은 품질을 보였으며, 기간에 관계없이 일정한 수준의 품질을 제공함을 알 수 있었다. 하지만 이러한 통합 프레임워크의 경우 데이터 적재를 수행하는 관리자가 많은 도메인 지식을 가지고 설정해야 하기 때문에 사용하기 어렵다는 단점이 있다. 상용 제품을 사용하는 경우, 2개월치 데이터에 대한 실험 결과는 1개월 데이터에 비해서 사용자가 설정하고 사용하기는 쉬운 반면, 사용자가 규칙을 어떻게 설정하느냐에 따라 많이 좌우되고 기준을 제시하지 않아 데이터 개수와 관계없이 매우 유동적인 결과를 보임을 알 수 있다.

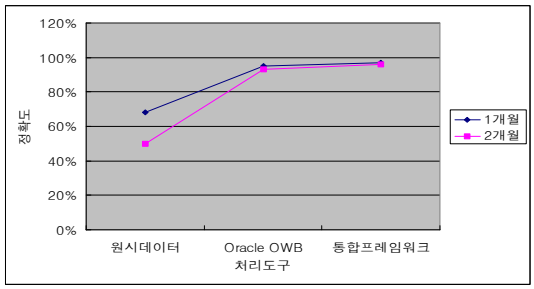

그림 7. 기간별 엔티티 품질평가 수행비교
Fig. 7. Experimental results of entity quality evaluation per data periods.

그림 8은 데이터 별로 엔티티에 대한 품질 평가를 비교한 내용이며, 원시데이터와 차원 데이터 그리고 원시와 차원 모두에 대해서 2개월 데이터 내에서 비교해 보았다. 그림에서 원시데이터는 검지기에서 취득되는 차량의 속도, 교통량 등의 계측 데이터이고 차원데이터는 고속도로 검지기 위치정보 등과 같이 계측데이터의 검색을 위해서 사용된 관련 정보 데이터이다. 그림에서 보는 바와 같이 통합 프레임워크 적용 시에는 데이터의 종류와 관계없이 그래프가 거의 수렴하는 것으로 보아 데이터가 일정한 수준의 품질을 제공하는 것을 알 수 있다. 또한, 상용 제품인 OWB를 사용하는 경우 현재 상용제품의 특징이 교통 데이터에만 적용될 수 있는 정제 기능을 제공하지 못하고 있어 데이터의 종류와 오류 처리에 대한 규칙 기술 정도에 따라 품질 결과는 많은 변화를 보임을 알 수 있다. 즉, 교통데이터 처리 프로세스를 완전하게 적용할 수는 없는 상태이다. Oracle OWB의 경우 원시와 차원의 경우가 차원데이터에 비해서 정확도가 높은 이유는 Oracle 데이터 OWB가 교통 데이터 차원의 특성을 잘 반영하지 못하기 때문이다.

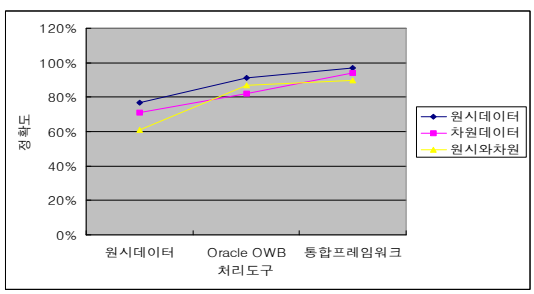

그림 8. 데이터별 엔티티 품질평가 수행비교
Fig. 8. Experimental results of entity quality evaluation per data values.

위와 같은 실험을 통해서 상용 오라클 OWB의 경우 교통데이터의 추출 및 적재를 위해 다양한 유형별로 실행하기 위한 편리함과 유용성을 제공하지만 좋은 품질을 보장하기 위해서는 개선의 여지가 있음을 알 수 있다.

Ⅴ. 결 론

ITS는 기존의 교통에 정보통신, 전기 및 전자, 기계, 제어 등의 기술을 접목시켜 도로 및 교통체계의 안전 및 운영 효율을 극대화하여 교통문제를 해결하는 종합적인 첨단 솔루션이다. 현재 국내에서는 ITS 사업 구축 시 실시간 교통데이터의 저장에 대한 활용방안을 고려하지 않은 채 단순히 주어진 집계간격과 데이터형식에 따라 저장하고 장기간에 걸친 데이터가 보관되지 않으므로 사용자가 원하는 고품질의 데이터를 추출 및 적재하고자 할 때 쉽게 데이터를 추출할 수 없는 큰 어려움이 있다.

본 논문에서는 이러한 요구에 따라 대량의 교통데이터웨어하우스를 구축하기 위한 시스템 구조와 상용 오라클 도구를 활용하여 운영계 시스템에서 교통데이터 웨어하우스로 교통데이터를 통합하기 위한 추출, 변환, 적재 방법을 제안하였으며, 효율적인 교통데이터의 품질관리를 위해서 데이터 값뿐만 아니라 데이터 구조 관점에서 통합 데이터평가 프로세스를 적용하여 실험 평가 하였다. 본 논문에서 실험 평가한 결과를 토대로 상용 ETL 제품인 OWB와 병행하여 사용하는 경우 ETL 효율뿐만 아니라 데이터품질 측면에서도 성능을 극대화 할 수 있을 것으로 기대된다.

감사의 글

본 논문은 2007년도 서일대학 학술연구비에 의해 연구되었음.

참 고 문 헌

[1] 한국도로공사 ITS사업실, *ITS 구축운영 편람*, 한국도로공사, 2005. 11.

[2] B. L. Smith, "An Investigation of ETL Techniques for Traffic Data Warehouse," TRB 2004 Annual Meeting.

[3] Smith, B., and S. Babiceanu, "Investigation of Extraction, Transformation, and Loading Techni- ques for Traffic Data Warehouses." *In Transpo- rtation Research Record 1879*, pp. 9-16, 2004.

[4] Smith, B. D. Lewis, R. Hammond. "Design of Archival Traffic Databases: Quantitative Investi- gation into Application of Advanced Data Mo- deling Concepts," *In Transportation Research Record 1836*, pp. 126-131, 2003.

[5] http://www.oracle.com/technology/documentation/warehouse.htm

[6] 박승용, *"고품질 데이터를 지원하는 교통데이터웨어하우스 구축 기법"*, 박사학위논문, 2007.

[7] W. H. Inmon, 통합 데이터베이스 구축 방법론, 홍릉과학출판사, 1997.

[8] R. Kimball, M. Ross, The Data Warehouse Toolkit, Wiley, 2002.

[9] Al-Deek, H. M., C. Chandra, "New Algorithms for Filtering and Imputation of Real-Time and Archived Dual-Loop Detector Data in I-4 Data Warehouse," *In Transportation Research Record 1867,* pp. 116-126, 2004.

[10] Wall, Z., and D. Dailey, "Algorithm for Detecting and Correcting Errors in Archived Traffic Data," *In Transportation Research Record 1855,* pp. 183-190, 2003.

우 찬 일 (禹讚溢)

1993년 2월 : 단국대학교 전자공학과(공학사)
1995년 2월 : 단국대학교 전자공학과(공학석사)
2003년 2월 : 단국대학교 전자공학과(공학박사)
2004년 3월~현재 : 서일대학 정보통신과 조교수
관심분야 : 정보보호 시스템, 데이터베이스 보안, 디지털 워터마킹

전 세 길 (全世拮)

1998년 2월 : 단국대학교 컴퓨터공학과(공학사)
2000년 2월 : 단국대학교 컴퓨터공학과(공학석사)
2004년 2월 : 단국대학교 전자컴퓨터공학과(공학박사)
2006년 1월~현재 : 한국도로공사 도로교통기술원 책임연구원
관심분야 : 정보보호 시스템, 데이터베이스 보안, 시공간 데이터베이스