

버스통계 데이터베이스의 합리적 구축 및 관리 방안 연구

A Study on the Design and the Management of Bus Database

손 상 훈

(경기개발연구원 연구원)

박 경 철

(경기개발연구원 책임연구원)

김 점 산

(경기개발연구원 책임연구원)

목 차

I. 서론

II. 이론적 고찰

1. 데이터베이스 구축방법론
2. 데이터베이스 품질관리

III. 버스통계자료 조사 및 분석

1. 버스통계자료 조사
2. 버스통계자료 분석

IV. 버스통계 데이터베이스 모형정립

1. 구축모형
2. 관리모형

V. 결론

참고문헌

1. 서론

대중교통의 한축을 담당하고 있는 버스서비스는 기저교통수단으로써 중요한 입지를 차지하고 있음과 동시에 지방자치 단체장의 주요 정책소재로 다루어져왔다. 그러나 이러한 중요성에도 불구하고, 버스 정책의 수립은 미시적이며, 단편적인 통계자료에 의존하여 추진되고 있는 실정이다. 하지만 최근의 흐름을 살펴보면, 교통정책에 대한 계획과 의사결정을 과학적, 체계적으로 수행하고 교통정책을 지속적으로 모니터링하기 위한 방편으로 객관적이며 보편적인 교통분야 기초자료로 구성된 데이터베이스 확보요구가 증대되고 있는 상황이다. 이러한 배경하에서 국가교통데이터베이스(KTDB)가 1990년대 후반부터 구축되기 시작하여 현재까지 국가단위의 교통정책 및 계획 수립 등에 필요한 다양한 기초자료를 수집/분석하여 수요자에게 제공하고 있으며, 일부 지자체에서도 지자체 단위의 교통정책 및 계획 수립을 지원하기 위해 교통데이터베이스 사업을 추진하고 있는 상황이다. 그러나 여전히 대중교통, 특히

버스분야에 대한 통계자료는 그 항목과 내용이 부족한 실정이며, 체계적인 수집 및 관리의 대상으로 자리잡고 있지 못하고 있다. 하지만 최근의 상황은 매우 긍정적이다. 버스자료와 관련한 전산화가 지속적으로 추진되고 있고, 교통카드가 상용화되고 버스 단말기가 설치됨에 따라 버스 운행정보와 요금정보가 실시간으로 수집/저장되고 있으며, 이러한 자료를 근간으로 활발한 연구가 진행되고 있다.

통신 및 전산기술의 발전은 인터넷의 활성화를 가져왔고, 통계정보 역시 다양한 방식의 콘텐츠로 제공됨에 따라 온라인상에서 누구나 접근하여 통계자료를 검색 후, 이를 분석/가공하여 정보로서 활용하고 있다. 대한민국의 대표적인 통계 데이터베이스인 통계청 통계정보시스템 KOSIS (Korean Statistical Information System)의 경우 2000년대 이후 급속히 사용빈도와 사용횟수가 증가한 것으로 보고되고 있다. 하지만 통계청(2000~2002) 조사 결과 사용자 만족지수는 오히려 2000년 75.9%에서 2년 만에 65.9%로 하락하는 상황이 발생하였는데, 이는 통계 데이터베이스가 양적으로 증대되고 있으나, 사용자층이 확대되면서 다양한 수요자

요구를 충족시키기 어려워지고 있으며, 지속적으로 사용자의 기대수준에 못 미치게 될 경우 많은 시간과 비용이 투자된 데이터베이스는 사용자로부터 외면당할 수도 있음을 경고하고 있다.

따라서 본 연구에서는 다양한 사용자 계층 및 요구수준을 충족시킬 수 있는 방안으로 기존의 기초현황 중심의 버스통계 자료의 수집과 제공 방식에서 탈피하여 계속적으로 수집되는 실시간 버스 운행 자료 및 버스승객의 통행패턴 자료를 활용한 새로운 형태의 통계 데이터베이스 구축 방안을 고찰하고, 데이터베이스 혹은 데이터에 대한 사용자의 부정적인 시각을 제거하기 위해서 품질관리 방안을 검토하고자 한다. 다만, 버스는 다양한 유형으로 구성되며, 유형에 따라 운행 특성 및 지역이 구분되고 있으므로, 본 연구에서는 유형적 범위로서 시내 버스, 지역적 범위로서 경기도 지역을 한정하여 광역권에 적절한 버스통계 데이터베이스 구축 및 관리 방안에 대한 검토를 수행하였다.

각 장의 내용은 다음과 같다. 2장에서는 데이터베이스 구축방법론 및 품질관리 이론에 대한 고찰을 실시하며, 3장에서는 합리적인 버스통계 데이터베이스 구축을 위해 통계자료의 수집체계와 통계자료목록을 조사/분석 한다. 4장에서는 구축모형, 관리모형을 제시하였고, 5장에서는 이를 바탕으로 결론과 향후연구과제를 제시하였다.

2. 이론적 고찰

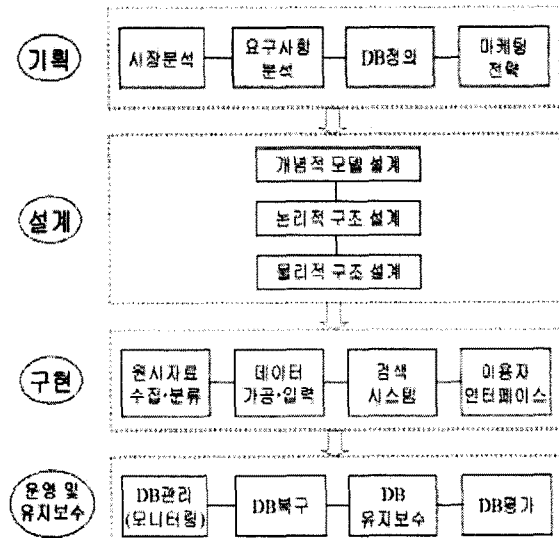
1. 데이터베이스 구축방법론

1) 데이터베이스의 정의

데이터베이스는 일반적으로 논리적으로 연관된 하나 이상의 자료의 모음으로, 그 내용을 고도로 구조화함으로써 검색과 갱신의 효율화를 꾀한 것이다. 즉, 몇 개의 자료 파일을 조직적으로 통합하여 자료 항목의 중복을 없애고 자료를 구조화하여 기억시켜 놓은 자료의 집합체이다. 때로는 데이터베이스를 자료의 제공범위까지 확장된 개념으로서 사용되기도 하는데, 본 연구에서는 후자의 관점으로 접근한다.

2) 데이터베이스의 구축과정

데이터베이스의 구축은 일반적으로 기획-설계-구현-운영/유지보수의 단계로 수행되는데, 기획단계에서는 데이터베이스 구축관련 제반사항을 규정하여 데이터베이스의 범위, 성격, 서비스를 명확하게 정의한다. 설계단계에서는 개체를 조건 및 관계를 통해 논리적으로 조직한다. 구현단계는 구체적으로 데이터베이스를 제작하는 단계로서, 데이터의 수집/가공/입력 및 소프트웨어의 설계와 설치가 이루어지는데, 적절한 데이터의 수집/가공은 데이터베이스의 품질과 직결되며, 이용자 만족도 확보의 밑거름이 된다. 또한 지속적이고 효율적인 데이터베이스 운용을 위해서는 운영 및 유지보수 단계가 반드시 필요하다.



<그림 1> 데이터베이스 구축과정(데이터베이스 구축 방법론, 한국데이터베이스진흥협회, 1998)

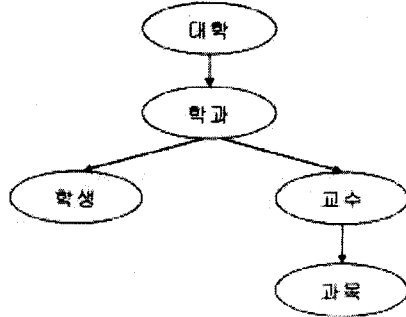
4) 데이터의 구조화

데이터를 구조화하기 위해서는 개체(entity), 속성(attribute), 관계(relationship)라는 개념을 사용한다. 개체는 데이터베이스에서 표현될 수 있는 사람, 장소, 사물, 개념, 사건과 같은 분명한 대상을 의미하며, 속성이란 객체의 어떤 관점을 설명하는 성질이고, 관계란 여러 객체간의 결합을 나타내는 개념이다.

(1) 계층형(hierarchical)

계층형 구조는 데이터가 상하 종속적인 관계

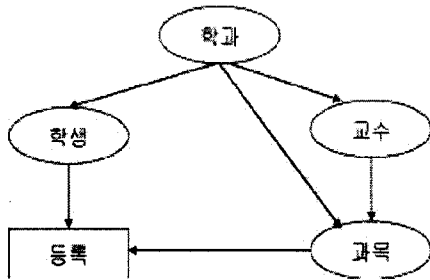
로 구성되어 있으며 데이터의 액세스 속도가 빠르고 정형화된 검색시스템에 유리한 구조이나, 데이터 관계의 수정이 어렵고, 변화하는 프로세서에 대한 적용이 어려운 것이 단점이다.



<그림 2> 계층형 구조

(2) 네트워크형(network)

네트워크형 구조는 데이터베이스의 구조를 네트워크(그래프) 형태로 나타낸 것으로 그래프 상의 노드가 레코드가 되며 이를 일대다 형태로 연결한 구조이다. 네트워크 구조는 레코드간의 어떠한 관계도 구현이 가능하여 큰 유연성을 가지지만, 구성과 설계가 복잡하고 궁극적으로 데이터의 종속성을 해결하지 못하며, 저장공간에 과부하가 발생하는 문제가 있다.



<그림 3> 네트워크형 구조

(3) 관계형(relational)

관계형 구조는 현재 가장 보편적으로 사용되는 구조로서, 데이터베이스를 구성하는 개체와 이들 개체간의 관계를 단순한 테이블 형태로 정의하였다. 이는 수학적 논리관계를 테이블 형태로 구성한 구조로서 테이블 내의 필드 중 일부를 다른 테이블과 중복하는 방법으로 각 테이블 간의 상관관계를 형성한다. 관계형 구조는 데이터의 생성과 갱신이 용이하여 조작성과 유연성이 뛰어나며, 높은 생산성, 응용 프로그램 개발이 쉽다는 장점이 있으나, 시스템의

부하가 상대적으로 높은 단점이 있다.

학생	학번	이름	주소	전공
학생(학번, 이름, 전공)				
과목	과목번호	과목이름	강사이름	
과목(과목번호, 과목이름, 강사이름)				
강사	강사이름	강사실		
강사(강사이름, 강사실)				
등록	학번	과목번호	성적	
등록(학번, 과목번호, 성적)				

<그림 4> 관계형 구조

(4) 객체지향형(object-oriented)

객체지향형 구조는 최근 급증하고 있는 멀티미디어 데이터의 원활한 처리를 위해 고안된 구조이며 객체 지향적 프로그래밍의 개념을 도입하여 현실세계에 나타나는 개체를 객체(object)로 정의하고, 객체의 상태에 따라 반응하는 행위(behavior)로서 정의된다. 그래픽, 비디오, 오디오와 같은 복잡한 데이터 유형을 단순하게 처리할 수 있고, 새로운 응용 및 적용이 매우 용이하나, 아직까지 개념과 기술에 대한 정립이 진행 중에 있다.

학생	클래스이름	과목
이름 생년월일 학번 주소 전화	속성 리스트	과목코드 과목이름 학점
calc-avg() calc-gpa() register-for(course)		연산 리스트
		enrollment()

<그림 5> 객체지향형 구조

2. 데이터베이스 품질관리

1) 데이터베이스의 품질

데이터베이스의 품질에 대한 정의는 데이터와 품질을 해석하는 사람마다 관점에 차이가 있으며 그 범위가 넓기 때문에 한마디로 단정하기 어렵다. 데이터베이스 구축 방법론(한국 데이터베이스진흥센터, 1998)에서는 <표 1>과 같이 데이터 품질, 서비스 품질, 기타 사항으로

구분된 품질기준을 통해 데이터베이스 품질의 속성을 정확성, 완전성, 최신성, 일관성, 검색성, 사용용이성, 이용자지원성, 비용, 네트워크/하드웨어, 일반적 요구사항으로서 정의하였다.

<표 2> 데이터베이스 품질기준

구분	품질기준	핵심관점
데이터 품질	정확성 (accuracy)	데이터 값의 정확성과 데이터 표현의 정확성
	완전성 (completeness)	데이터 구조의 완전성 데이터 값의 완전성 데이터 표현의 완전성
	최신성 (currentness)	데이터의 갱신 및 낙후 정도
	일관성 (consistency)	데이터베이스 데이터간 불일치 정도
서비스 품질	검색성 (searching)	검색과정에서 제공되는 방법, 도구, 매체들의 적합정도
	사용용이성 (easy of use)	인터페이스 및 유틸리티 기능의 편리성, 친절성, 이해가능성 등
	이용자지원성 (customer support)	이용자에게 제공되는 교육 훈련, 각종문서 도움 서비스의 적합정도
	비용 (cost)	소요비용 데이터베이스이용 비용
	네트워크/하드웨어 (network hardware)	데이터베이스의 이용 속도
기타	일반적요구사항 (general fact desired)	일반적 요구사항

자료: 데이터베이스 구축 방법론, 한국데이터베이스진흥센터, 1998

2) 품질관리

품질관리란 기관이나 조직 내외부의 정보시스템 및 데이터베이스 사용자의 기대를 만족시키기 위해 지속적으로 수행하는 데이터베이스 관리 및 개선활동을 의미한다. 최근 정보시스템이 보편화되면서 각 부분별 데이터 간에 심각한 중복성과 불일치성의 문제가 제기되고 있는 상황이며, 각 기관의 내부 시스템 통합뿐만 아니라 외부 시스템과의 연계가 필요하게 되면서 데이터베이스의 품질관리가 중요하게 되었다.

(1) 품질관리 프로세스

데이터베이스 품질관리 프로세스는 8개로 구성되며, 각 프로세스별 상세내용은 <표 2>와 같다.

<표 3> 데이터베이스 품질관리 프로세스

프로세스	내용
요구사항 관리	데이터에 대한 사용자의 요구를 도출/관리하는 과정
데이터표준 관리	데이터에 대한 공통된 시각을 유지할 수 있도록 데이터 표준화 원칙을 정의하고 통제하는 과정
데이터오너십 관리	데이터에 책임과 권한을 가진 오너를 중심으로 품질관리 절차를 정의하고 실행하는 과정
데이터구조 관리	사용자 요구 데이터를 제공할 수 있도록 데이터 모델 및 데이터베이스를 설계/관리하는 과정
데이터베이스 관리	유동적이고 변화가 많은 데이터베이스에 대한 상황에 따른 대응으로서 성능, 보안 관리 과정
데이터흐름 관리	원시데이터를 가공, 추출, 변환, 적재하는 제반과정을 관리하는 과정
데이터활용 관리	사용자의 데이터 활용 현황을 파악하고 이를 극대화할 수 있는 환경을 구현하는 과정
사용자 뷰 관리	사용자가 최종적으로 데이터를 이용하는 환경이 사용자 관점에서 설계·구현되도록 관리

자료: 데이터 품질관리 성숙모형, 한국데이터베이스진흥센터, 2006

(2) 품질관리 성숙모형

품질관리 성숙모형은 조직이 수행하고 있는 품질관리의 체계화 정도를 판단하는 기준이 되며, 품질수준 측정 및 품질 향상을 위한 방향을 제시한다. 데이터 품질관리 성숙모형(데이터 품질관리 성숙모형, 한국데이터베이스진흥센터, 2006)은 데이터 품질관리 성숙수준을 도입-정형화-통합화-정량화-최적화로 구분하였다.

<표 4> 데이터 품질관리 성숙수준

최적화	데이터 품질관리 개선에 필요한 요소를 지속적으로 도출/적용
정량화	정량적인 측정방법을 통해 데이터 품질관리가 수행
통합화	전사적인 연계 통합과정에서 일관성있는 데이터의 품질 관리가 수행
정형화	데이터 품질관리를 위한 프로세스가 정형화된 단계
도입	데이터 품질관리의 문제점과 필요성에 대해 인식하고 있으며 부분적인 데이터 품질관리 활동을 시행

3) 품질평가

데이터베이스 표준화 연구의 일부로서 수행된 데이터베이스 품질평가 항목(한국데이터베이스진흥센터, 2000)에서는 품질평가의 기준과 평가항목을 제시하였고, 측정방법, 도구, 절차를 제안하였다. 평가기준과 평가항목에 의해 데이터베이스의 품질을 측정하여 우수한 부분과 미흡한 부분을 선별하고, 이를 통해 품질보완 및 개선을 수행한다.

(1) 품질평가 절차

품질평가의 첫 단계로서 측정대상 데이터베이스를 분석하고 평가목적에 따른 측정범위를 선정하여 그에 따른 평가항목을 구체화한다. 측정범위는 데이터 질, 데이터베이스 서비스 혹은 양자 모두 중 하나로 결정된다. 두 번째 단계로서 응답자 선정이 요구된다. 응답자 선정이 객관적이며, 공정하게 이루어질 수 있도록 주의한다. 세 번째 단계에서 실제조사가 이루어진다. 조사 방법으로서 E-mail, 우편, 전화, 방문 등의 방법이 있으며, 조사를 통해 평가항목별로 채점을 실시한다. 네 번째로 품질지수를 산정한다. 품질지수는 평가항목별 중요도와 항목별 품질기준에 대한 평균 응답을 곱하여 최종 품질지수를 산출하는데, 평가항목별 중요도는 네 번째 단계에서 응답자에게 개별적, 즉각적으로 품질기준의 중요도를 설문하거나, 미리 데이터베이스를 주제별, 가공형태별, 표현형태별, 언어별, 제공매체별 등으로 분류하여 각 분류별로 데이터베이스의 품질 기준 중요도를 산출한 후 적용할 수 있다.

1단계	- 측정대상 데이터베이스 분석 - 평가목적에 따른 측정범위 선정
2단계	- 응답자 선정
3단계	- 실제 조사 (평가항목별 채점)
4단계	- 품질지수 산정

<그림 6> 품질평가 절차

(2) 품질등급 및 결과

품질등급제를 통해 데이터베이스를 차별적으로 등급화하고, 각 등급에 요구되는 품질수준을

명시함으로써 데이터베이스 품질에 대한 사용자 인식을 높이고 개발자 노력을 유도할 수 있다.

<표 5> 품질등급별 품질수준

등급명	내용	해당 품질지수
1등급	매우만족	75점이상
2등급	만족	50점~74점
3등급	보통	25점~49점
4등급	만족못함	25점 미만

3. 버스통계자료 조사 및 분석

1. 버스통계자료 조사

1) 버스통계자료의 정의

버스통계자료란 버스와 관련하여 실시간으로 혹은 일정시점에 대하여 혹은 일정기간동안 수집하여 수량화한 자료를 의미한다. 예를 들어 버스운행 현황은 실시간 자료, 2006년 12월 31일 기준 버스업체수는 일정시점의 자료, 2006년 버스사고 수는 일정기간의 자료가 된다.

2) 수집/관리/제공 주체 조사

버스통계 데이터베이스를 구축하기 위해서는 현재 다양한 기관에서 다양한 형태로 제공되고 있는 버스통계자료에 대한 수집이 선행되어야 한다. 이를 위해 먼저 버스통계자료를 수집/관리/제공하고 있는 대상을 파악하였고, 각각의 특징은 <표 5>와 같이 정리하였다.

<표 6> 버스통계 수집/관리/제공 주체

대상	특징 / 주요 버스관련 자료
통계청 통계정보시스템 (KOSIS)	국내의 주요통계정보를 수록한 대표적 통계DB로, 운수산업 관점에서의 버스자료 제공 (사업체수, 종사자수, 매출액 등)
국가교통 데이터베이스 (KTDB)	국가단위의 교통정책/계획 수립에 필요한 기초자료 및 통계를 조사/수집 (분석원시자료, 통계청, 건교부 제공 버스자료 등)
경기도교통 데이터베이스 (구축중)	경기도 차원의 교통데이터베이스로서 교통전반의 자료를 수집 후 가공/분석하여 제공 (경기도권역 분석 원시자료, 버스운행 가공자료 등)

경기도(지자체) 버스운행관리/ 정보시스템 (BMS/BIS)	버스의 운행상황을 실시간으로 파악하고 시민들에게 버스 운행 정보를 안내하여 운행관리 및 이용자편의 극대화 추구 (버스위치, 버스속도, 승하차인원 등)
전국운송 사업조합 버스통계편람	버스운송사업 발전과 이익을 도모하기 위한 조직. 버스 운송사업의 통계작성 및 조사수행 (운송, 차량, 종사자, 노선 통계 등)
경영 및 서비스 평가자료	버스업체 재정지원을 위한 평가 기준자료로서 전업체의 일반, 재무, 서비스현황 자료가 조사/ 분석됨 (인원, 차량, 노선, 수입금 등)
행정기관 (버스업무편람, 버스인면허 대장 외)	버스의 인허가 및 버스관련 시설 업무를 담당하는 기관으로서 전체적인 버스 현황 및 재정, 법률자료 보유 (노선, 시설 내역 및 현황 등)
경찰청/ 도로교통 안전관리공단/ 손해보험사	버스사고와 관련한 자료들이 수집/가공/제공되고 있음. (사고위치, 규모, 피해정도 등)

<표 7> 버스통계자료 상세목록 및 그룹화

그룹	상세목록
버스수송 실적자료	수송인, 수송인-km, 수단분담률, 환승통행량 등
버스업체 현황자료	업체명, 소속시군, 사업유형, 종사자수, 등록차량 및 노선 등
버스차량 현황자료	연료별 구분, 차량, 검사정보 등
버스노선 현황자료	유형별 구분, 등록차량수, 기종점, 배차간격, 운행거리 등
버스운행 자료	버스위치, 속도, 배차준수율, 정류장별 승객수 등
버스시설 자료	버스정류소, 버스터미널, 버스전용차로, 공영차고지 등
버스요금 자료	버스유형별 요금, 카드이용률
버스경영실태 자료	노선별 수입금, 인건비, 연료비, 경상수지 등
버스서비스수준 현황	이용자만족도, 차내서비스 결과 등
버스사고 자료	사고일시, 사고위치, 피해정도, 대인/대물 유형, 과실율 등
버스행정처분 현황	민원접수기관, 민원내용, 처리결과 등
버스재정지원 내역	재정지원 유형, 지원액 등
버스관계 법령자료	법, 법령, 조례, 훈령명, 제정 및 개정일, 개정사유 등
버스관련기관 현황	기관명, 유형, 주요업무, 연락처, 홈페이지 등
버스관련 연구자료	연구명, 연구기간, 발주/수행처

3) 상세 목록 조사

전체 버스통계자료 목록을 수집하여 <표 6> 과 같이 15개의 주제별로 그룹화 하였다

2. 버스통계자료 분석

1) 수집/제공체계 분석

(1) 시간적 분석

시간적 분석으로서 현재 버스통계자료 수집 기관의 수집주기의 유형에 대한 분석을 실시하였다. 수집주기는 주기적, 비주기적으로 구분하였는데, 주기적 수집의 경우 BMS/BIS를 제외한 나머지의 경우는 통상 1년 주기로 자료를 수집하며, BMS/BIS의 경우 실시간으로 계속해서 자료가 수집된다.

<표 8> 시간적 분석 결과

유형	대 상	계
주기적	KOSIS, 국가교통DB, 경기도교통DB, 운송사업조합, 안전공단, 경영/서비스평가, BMS/BIS	7
비주기적	행정기관, 경찰청, 손해보험사	3

(2) 공간적 분석

공간적 분석으로서 현재 버스통계자료 수집 기관의 수집범위에 대한 분석을 실시하였다. 수집범위는 거시적, 미시적으로 구분하였으며, 거시적일수록 지자체, 광역지자체, 국가로 그 범위가 확대되고, 미시적일수록 개별 업체, 노선, 사고 등으로 구체적인 대상으로 범위가 좁혀진다.

<표 9> 공간적 분석 결과

범위	대 상	계
거시	KOSIS, 국가교통DB, 운송사업조합, 안전공단	4
	경기도교통 DB	1
미시	BMS/BIS, 경영/서비스평가, 행정기관, 경찰청, 손해보험사	5

(3) 수집형태 분석

수집형태 분석은 버스 통계자료 수집형태를 원시자료 수집과 가공자료 수집으로 구분한다.

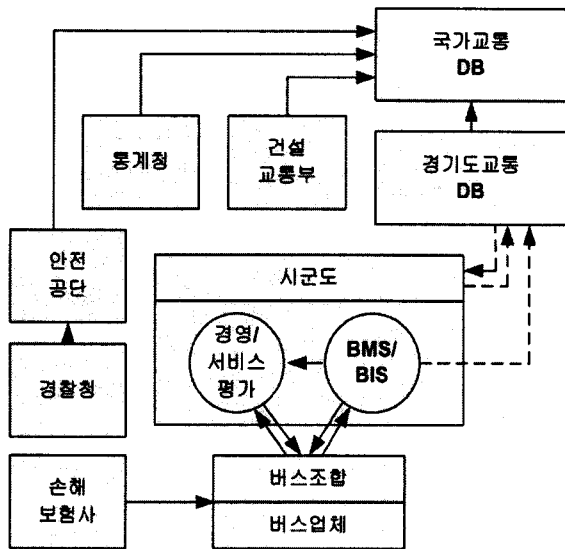
원시자료 수집은 원시자료를 직접 수집/제공하는 형태이며, 가공자료 수집은 원시자료 수집을 거쳐 1차 가공된 자료를 재수집하는 형태이다.

<표 10> 수집형태 분석 결과

형태	대상	계
원시자료 수집	KOSIS, 운송사업조합, 경영/서비스평가, BMS/BIS, 행정기관, 경찰청, 손해보험사	7
가공자료 수집	국가교통DB, 경기도교통DB, 안전공단	3

(4) 수집/제공체계 다이어그램

버스통계자료의 이동 흐름을 <그림 7>과 같이 정리하였으며, 우측상단으로 갈수록 거시적, 가공자료 수집특성을 가지며, 하단으로 갈수록 미시적, 원시자료 수집특성을 지닌다.



<그림 7> 버스통계자료 이동흐름

2) 버스통계자료 분석

(1) 중복도 분석

중복도 분석은 각각의 버스통계자료가 얼마나 중복되어 수집되는가에 대한 분석으로서 자료수집 체계의 비효율성을 평가하는 척도로 사용될 수 있다. 앞서 구분한 것과 같이 원시자료를 직접 수집하는 대상에 한해서 <표 11>과 같이 버스통계자료 수집기관별 수집자료 목록집계를 통해 중복도를 분석한 결과 버스업체, 버스차량에 대한 자료는 각각 5곳, 4곳에서 수

집하고 있는 것으로 나타났으며, 버스노선과 버스경영관련 자료도 3곳에서 수집하고 있는 것으로 분석되었다.

(2) 질적특성 분석

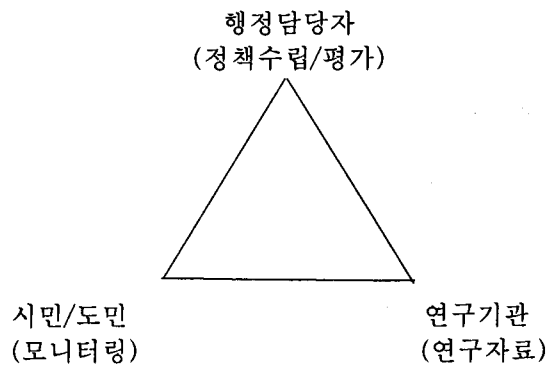
버스통계자료의 질적 특성을 이력자료, 현황자료, 실적자료로서 구분하여 분석하였다. 각각의 특성은 <표 10>과 같다. 이러한 구분을 통해 자료의 성격을 규정할 수 있는데, 이력자료는 원시자료(raw data), 현시자료는 자료(data), 실적자료는 정보(information)로 각각 대변할 수 있다. 원시자료-자료-정보의 순서로 그 가치는 증가한다. 수집된 버스통계자료 목록에 대한 질적특성 분석 결과는 <표 12>와 같다.

<표 11> 질적특성 분석

구분	특성
이력자료 (historical)	시간순서대로 빠짐없이 기록/수집한 자료 (인면허사항, 사고조서 등)
현시자료 (current)	매년 말 혹은 매월 말과 같이 특정시점을 기준으로 규모/현황을 나타낸 자료 (버스업체수, 버스노선수 등)
실적자료 (periodical)	일년 혹은 분기와 같이 일정기간 동안 변화내역 및 실적을 집계한 자료 (년간 교통사고수, 분기별 수송승객 수 등)

3) 이용자 분석

버스통계 데이터베이스를 활용하는 사용자 그룹을 행정담당자, 시민/도민, 연구기관으로 구분하였고, 각 사용자 그룹에 대한 특성을 <그림 8>과 같이 분석하였다.



<그림 8> 이용자 분석

<표 12> 버스통계자료 수집기관별 수집자료 목록

구분	버스수송	버스업체	버스차량	버스노선	버스운행	버스시설	버스요금	버스경영	버스서비스	버스사고	행정처분	재정지원	버스법령	버스기관	연구자료
통계정보시스템	✓	✓						✓							
운송사업조합	✓	✓	✓				✓	✓							
행정기관(인면허대장)		✓	✓	✓		✓	✓				✓	✓	✓		✓
경영 및 서비스평가		✓	✓	✓				✓	✓	✓	✓				
BMS/BIS		✓	✓	✓	✓	✓	✓								
경찰청, 손해보험사										✓					
중복도	2	5	4	3	1	2	3	3	1	2	2	1	1	0	1

<표 13> 자료의 질적특성 분석 결과

구분	버스수송	버스업체	버스차량	버스노선	버스운행	버스시설	버스요금	버스경영	버스서비스	버스사고	행정처분	재정지원	버스법령	버스기관	연구자료	계
이력자료 (historical)		✓	✓	✓		✓	✓			✓	✓	✓	✓		✓	10
현시자료 (current)		✓	✓	✓	✓	✓	✓						✓	✓		8
실적자료 (periodical)	✓				✓			✓	✓	✓	✓	✓			✓	8

4. 버스통계 데이터베이스 모형정립

1) 구축모형

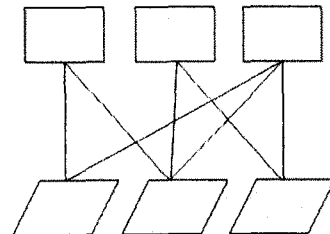
(1) 기획

단순히 버스와 관련하여 수집된 원시데이터를 축적하여 제공하는 매개체로서의 버스통계 데이터베이스가 아니라 이용자의 요구에 맞는 정보를 제공하고, 이용자가 원하는 정보를 쉽게 얻도록 하는 고차원적인 역할을 수행할 수 있는 버스통계 데이터베이스 구축을 목표로 한다. 다만, 본 연구에서는 국가단위가 아니라 광역지자체 규모에 적합한 데이터베이스로서 정책수립을 위한 장기적/거시적 정보와 함께 버스운행에 대한 관리가 가능하도록 단기적/미시적인 정보까지 제공 가능한 데이터베이스이어야 한다.

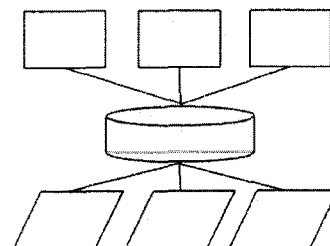
(2) 설계

앞서 중복도 분석에서 나타난 결과와 같이 동일한 통계자료에 대하여 여러기관에서 중복하여 수집하고 있는 것은 비효율성을 의미한다. 한 번의 수집으로 여러기관에서 활용할 수 있는 체계를 구축해야 하며, 따라서 기존의 수집체계를 잘 활용하여 정보의 허브(hub) 역할을 수행할 수 있는 형태로 데이터베이스를 구축한다. 수집주기는 현재 구축된 전산인프라를

고려할 때 한 달로 해도 무리가 없다 판단된다. 자료의 수집/가공 체계는 <그림 9>와 같이 각각의 자료가 여러 수집기관에 의해 중복 수집되는 것에서 <그림 10>과 같이 버스통계 데이터베이스에 의해 수집된 자료는 다른 기관에 제공할 수 있도록 한다. 이력자료는 최초 수집기관에 저장/보관되고, 실적자료 및 분석자료 경우 버스통계 데이터베이스에서 주기적으로 수집하는 체계가 보다 효율적이라 판단한다.



<그림 9> 기존 수집/가공체계



<그림 10> 제안 수집/가공체계

(3) 구현

데이터의 가공과 표출에 있어 현황자료와 실적자료의 결합을 시도하였다. <표 13>은 버스노선의 현황자료이다. 이는 버스인면허의 가장 최신자료로서 월마다 정기적으로 정리하여 버스통계 데이터베이스로 송부한다. <표 14>는 버스노선별 BMS/BIS 운행자료를 통해 가공이 가능한 자료로서 이를 한 달 주기의 실적자료로 가공하여 버스통계데이터베이스로 송부한다. 버스통계 데이터베이스에서는 이들 자료를 결합하여 사용자의 요구수준에 맞게 표출하는 역할을 수행한다.

2) 관리모형

(1) 품질관리

데이터베이스에 있어 품질관리의 중요성은 앞서 언급하였듯이 데이터베이스의 이용자 만족도 증진 및 오정보로부터 유발되는 피해를 방지하기 위해 반드시 필요하다. 버스통계 데이터베이스의 품질관리 대상은 데이터와 데이터서비스를 포함하는 데이터베이스 전체이다.

품질관리는 품질관리 성숙모형을 통해 수행한다. 제1단계로서 품질관리에 대한 중요성을 인식하여 버스통계 데이터베이스의 일부로서 품질관리방안을 반드시 포함하며, 제2단계에서는 프로세스를 구체화하여, 제3단계에서는 내부적으로 품질관리를 수행한다. 제4단계에서는 평가를 통한 품질지수를 정량화하고, 제5단계에서는 이를 지속적으로 수행할 수 있도록 추진한다.

<표 14> 버스노선 현황자료(예시)

노선번호	유형구분	업체명	등록대수(대)	운행횟수(회/일)	경유정류장(개소)	기종점/경유지 목록	첫차시간		배차간격(분)			막차시간		운행거리(km)
							주중	주말	첨두	비첨두	주말	주중	주말	
000														
...														

<표 15> 버스노선 실적자료(예시)

노선번호	첫차시간 준수율(%)	배차간격 준수율(%)	막차시간 준수율(%)	평균통행속도(km/h)		평균수송인원(인/시간)	
				주중(0시~24시)	주말(0시~24시)	주중(0시~24시)	주말(0시~24시)
000							
...							

(2) 평가

성숙모형의 제4, 5단계인 최적화와 정량화를 달성하기 위한 방안으로 데이터베이스에 대한 평가를 주기적으로 수행한다. 일반적으로 데이터베이스 평가는 데이터베이스 구축 전단계에 걸쳐서 지속적이고 반복적으로 행해지지만 특히 데이터베이스의 사후평가는 운영되는 데이터베이스의 품질, 가치, 효과를 측정하여 판단하고 이에 따른 문제점을 파악하여 이를 해결하고 개선하는 것을 목표로 한다.

데이터베이스의 무엇을 평가할 것인지를 정확히 파악하기 위해서는 데이터베이스 가치발생 단계에 따른 평가 대상을 파악해 볼 필요가 있다. 이는 시간이 지남에 따라 가치가 조금씩 다른 형태를 취하게 되고, 가치의 형태에 따라 평가주체, 기준, 대상, 기법이 각기 다르게 구성되어야 하기 때문이다.

평가기준은 정확성, 완전성, 최신성, 일관성, 검색성, 사용용이성, 이용자지원성, 비용, 네트워크/하드웨어, 기타 일반적 요구사항을 중심으로 하되 항목의 선정 및 항목별 중요도의 선정은 버스통계 데이터베이스 사용자 그룹에 의해 정하도록 한다.

5. 결론

본 연구는 효율적이고 합리적인 버스통계 데이터베이스 모형 도출을 목표로 하여 이를 위한 데이터베이스 구축방법론과 품질관리 제반 이론을 검토한 후, 버스통계자료의 수집체계

및 상세목록 분석을 통해 광역지자체에서 활용 가능한 버스통계 데이터베이스 모형을 제시하였다. 모형에서는 목표정립, 수집기관분석, 통계목록분석, 사용자분석, 데이터베이스 구조, 자료형태 및 품질관리/평가를 포함하고 있다. 이를 통해 다양한 사용자 요구에 부응하고, 정책도출, 버스산업 및 운행 모니터링, 버스관련 연구자료 제공이 가능할 것으로 판단된다. 향후연구과제로서 버스 혹은 교통통계 데이터베이스의 품질평가를 위해 버스 혹은 교통통계 데이터베이스에 적합한 품질기준 및 중요도의 선정을 통한 사후평가를 제시하며, 아울러 행정 전산화의 진행에 따른 버스통계 데이터베이스 모형의 수정연구가 지속적으로 요구된다.

참고문헌

- 1) 김선호, 데이터베이스시스템 모델링과 디자인, 삼인커뮤니케이션, 2001.
- 2) 경기개발연구원, 경기도 교통데이터베이스 구축 및 활용방안에 관한 연구, 2005.
- 3) 경기도, 경기도 버스운송업체 경영 및 서비스 평가, 2006
- 4) 이점남, 사용자 관점에서의 KOSIS 통계 DB 품질평가 연구, 한남대학교, 2003
- 5) 전국버스운송사업조합, 버스통계편람, 2006.
- 6) 정보통신부, 데이터베이스 구축 방법론, 한국데이터베이스진흥협회, 1998.
- 7) 정보통신부, 데이터베이스 품질평가 항목, 한국데이터베이스진흥센터, 2000.
- 8) 정보통신부, 데이터 품질관리 성숙모형(Ver 1.0), 한국데이터베이스진흥센터, 2006.
- 9) 정보통신부, 데이터 품질관리 지침(Ver 2.1), 한국데이터베이스진흥센터, 2006.
- 10) 조중래, 교통DB, 그 10년을 생각하며, 교통기술과 정책, 제2권 제4호, 2005.12
- 11) 최기주, 교통데이터와 교통정책(Transportation Data and Policy): 연결파리가 시급하다, 교통기술과 정책, 제2권 제4호, 2005.12
- 12) 통계청, 행정서비스에 대한 고객만족도 조사 결과보고서, 현대리서치연구소, 2000~2002.
- 13) 경기도교통DB센터, 경기도내부자료
- 14) 경기도버스은행관리시스템, 경기도내부자료
- 15) 국가교통DB센터(<http://www.ktdb.go.kr/>)
- 16) 경기도 버스정보시스템(<http://www.gbbs.go.kr/>)