

JSON을 이용한 고객 물류 데이터 통합 시스템*

장현애, 정희진, 이은지, 최다은, 박혁주, 송양의, 이용규
동국대학교 컴퓨터공학과-서울
e-mail : haleyjang1102@gmail.com

A System for Integration of Customer Distribution Data Using JSON

Hyeon Ae Jang, Hee Jin Jeong, Eun Ji Lee, Da Eun Choi, Hyeok Ju Park,
Yang-Eui Song, Yong Kyu Lee
Department of Computer Science and Engineering, Dongguk University-Seoul

요 약

소규모 전자상거래 업체에서 데이터베이스를 사용하여 고객 물류 정보를 관리하지 못하고 파일로 관리한다. 파일로 관리되는 데이터는 자료 형식의 차이로 업체 간 데이터를 주고받기 어려운 문제가 있다. 이러한 문제를 해결하기 위해서 파일로 관리되는 고객 물류 데이터를 하나의 시스템으로 통합하여 관리할 수 있는 체계가 필요하다. 본 논문에서는 JSON을 이용한 고객 물류 관리 시스템을 제안하여 통합 문제를 해결하고자 한다. 본 시스템은 각 고객 물류 데이터의 데이터 구조를 분석하여 통합 데이터베이스를 위한 스키마를 생성하고, 새로운 스키마에 맞게 기존 데이터를 전처리 과정을 거쳐 통합 데이터베이스에 저장하여 관리한다. 제안된 시스템을 사용하면 파일로 입력된 데이터가 자동으로 통합 데이터베이스로 저장되어, 업체별로 관리되는 데이터를 통합하여 관리할 수 있다.

1. 서론

다양한 분야에서 생산되는 데이터의 양이 증가되면서 체계적인 관리를 위해 데이터베이스를 사용한다. 전자상거래 분야에서도 고객의 정보와 주문정보를 관리하기 위해서 데이터베이스를 사용하고 있다. 업체별로 데이터베이스를 구축하기 때문에, 업체의 업무 특성에 따라 각기 다른 스키마와 데이터 형태로 데이터베이스를 설계한다[1]. 이러한 경우 중간 유통 관리자는 물류의 흐름을 파악하기 위해서 업체별로 관리되는 시스템에 접근하여 내용을 파악하기 때문에 업무처리 시간이 오래 걸리고 비효율적으로 처리되게 된다. 이 때 발생하는 문제점을 보완하기 위해서 데이터베이스 통합 시스템이 필요하다.

대규모 물류관리 기업에서는 전사적자원관리(Enterprise Resource Planning, ERP)를 이용하여 기업 내 생산, 물류, 재무, 회계, 영업과 구매, 재고 등 경영활동 프로세스를 통합적으로 연계해 관리하고 있다. 하지만 소규모의 기업에서는 비용 부담으로 ERP 시스템을 도입하기에 한계가 있어 회사 운영에 필요한 고객물류 데이터를 파일에 수작업으로 관리하는 경우가 많다[2]. 파일로 관리되는 정보는 여러 사람이 작업하거나 여러 업체의 정보를 체계적으로 관리하기 어려운 문제가 있다.

본 논문에서는 데이터를 파일로 관리하여 업무를 처리

하는 데에서 오는 비효율성과 생산성 저하의 문제를 해결하기 위해 JSON을 이용한 고객 물류 데이터 통합 시스템을 제안하고자 한다. 본 논문에서 제안하는 시스템에서는 전자상거래 업체에서 관리하는 고객물류 데이터를 분석하여 통합된 데이터베이스 스키마를 설계하고, 고객물류 데이터 파일을 JSON로 변환한다. 변환된 JSON은 통합 데이터베이스에 저장하여 관리한다[3]. 통합된 데이터베이스를 사용하여 통일성과 효율성을 보장하고, 통일된 규격의 데이터베이스 사용이 가능하도록 한다.

이 논문의 구성은 다음과 같다. 2장에서는 관련연구로 데이터베이스 통합과 JSON을 소개하고, 3장에서는 제안한 고객 물류 데이터 통합 시스템의 구조와 각 모듈별 기능을 설명하고, 4장에서는 구현결과와 통합 데이터베이스의 스키마를 소개한다. 마지막 5장에서는 결론을 맺는다.

2. 관련 연구

2.1 데이터베이스 통합

데이터베이스 통합은 정보처리 분야의 증가하는 복잡성을 개선할 수 있는 기술로, 다수의 시스템 상에서 분산 운영되고 있는 자료의 중복 제거 및 비호환성 문제를 해결하기 위해 사용한다.

각 기업에서 각기 다른 데이터베이스를 운영함으로써 기업들 사이의 데이터베이스가 상호 호환을 이루지 못하여 데이터베이스 통합하기 위한 노력이 이어지고 있다. 기존의 연구 방법은 데이터베이스 통합을 위해 많은 시간과

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원사업의 연구결과로 수행되었음(2016-0-00017)

비용을 필요하기 때문에, 저비용으로 데이터 통합을 할 수 있는 방법이 요구된다[4][6].

2.2 JSON(JavaScript Object Notation)

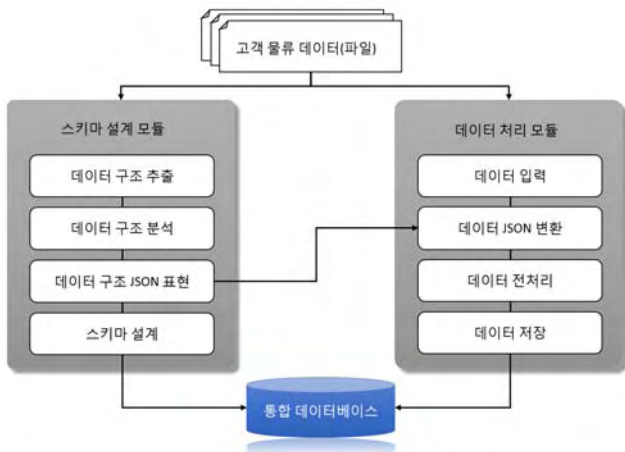
JSON은 프로그래밍 언어에 독립적인 텍스트 기반의 데이터 교환 형식이다. 이 형식은 ECMAScript 프로그래밍 언어 표준에 토대를 두고 있으며 key/value 형태의 쌍으로 이루어진 구조적 데이터의 표현을 정의한다[5].

JSON은 프로그래밍 언어와 플랫폼에 독립적이므로 C, C++, C#, Java, JavaScript, Perl, Python 등 대부분의 프로그래밍 언어를 지원하여, 서로 다른 시스템 간에 데이터를 교환하기에 적합하다. 또한 문법 규칙이 엄격한 XML을 사용할 때보다 수월하고 가볍고 손쉽게 데이터를 처리할 수 있다는 장점이 있다.

3. 고객 물류 데이터 통합 시스템

3.1 시스템 구조

(그림 1)은 본 논문에서 제안하는 시스템의 구성을 도식화한 것이다.



(그림 1) 고객 물류 데이터 통합 시스템 구조

고객 물류 데이터 통합 시스템은 스키마 설계 모듈과 데이터 처리 모듈로 구분된다. 고객 물류 데이터는 파일 형태이다. 스키마 설계 모듈은 데이터 구조 추출, 데이터 구조 분석, 데이터 구조 JSON 표현, 스키마 설계 단계로 진행된다. 데이터 처리 모듈은 데이터 입력, 데이터 JSON 변환, 데이터 전처리, 데이터 저장의 단계로 진행된다. 스키마 설계 모듈은 고객 물류 데이터의 데이터 구조를 분석하여 통합 데이터베이스의 스키마를 설계하고, 데이터 처리 모듈에서 고객 물류 데이터를 입력 받아 통합 데이터베이스에 저장한다.

3.2 스키마 설계 모듈

스키마 설계 모듈에서 데이터 구조 추출 단계는 여러

업체의 고객 물류 데이터 파일에서 애트리뷰트를 추출한다. 데이터 구조 분석 단계는 추출된 애트리뷰트명은 다르지만 동일한 의미를 가지는 애트리뷰트를 공통된 애트리뷰트로 정의된다. 데이터 구조 JSON 표현 단계는 공통 애트리뷰트를 기준으로 업체명을 key, 애트리뷰트를 value로 표현한다. (그림 2)는 생성된 JSON 형식으로 표현한 결과이다. 생성된 JSON은 스키마 설계 단계와 데이터 JSON 변환 단계에서 입력으로 사용된다. 마지막으로 스키마 설계는 각 JSON 파일에서 공통으로 갖는 애트리뷰트로 구성되어 있는 통합 테이블과 업체별로 고유한 애트리뷰트로 구성되어 있는 테이블로 데이터베이스 스키마를 설계한다.

```
{
  "수취인명": [
    {
      "A기업": "받는분 이름",
      "B기업": "받는 사람",
      "C기업": "수취인 이름"
    }
  ],
}
```

(그림 2) 데이터 구조 JSON 표현 결과

3.3 데이터 처리 모듈

데이터 처리 모듈의 데이터 입력 단계는 고객 물류 데이터 파일에서 고객의 상품 주문정보를 입력받는다. 데이터 JSON 변환 단계는 입력된 데이터를 스키마 설계 모듈에서 생성된 결과를 key로 설정하고, 데이터를 value로 하여 JSON 형식으로 변환한다. (그림 3)는 입력받은 고객 물류 데이터 파일을 JSON 형식으로 변환한 결과이다. 변환된 JSON은 데이터 전처리에서 데이터의 통합 데이터베이스의 각 업체별로 상이하게 구성되어 있는 데이터 형식을 일관된 형식으로 전처리 과정을 거치고, 마지막으로 데이터 저장 단계에서 통합 데이터베이스에 저장된다.

```
{
  "받는분 이름": "홍길동",
  "우편번호": "123-456",
  "주소": "서울특별시 중구 필동로 1길30",
  "받는분 휴대폰1": "010-1234-5678",
  "받는분 휴대폰2": "02-123-4567",
  "업체명": "가나다",
  "주문번호": "012345678",
  "구매수량": "2",
  "옵션명": "*한정특가*고구마 샐러드1kg",
  "배송메시지": "없음",
  "상품명": "[97무배] 샐러드 대용량 1KG",
  "구매자이름": "홍길동",
  "합포장번호": "987654321",
  "배송비유형": "선결제"
}
```

(그림 3) 물류 데이터 JSON 변환 결과

데이터 처리 모듈의 데이터 전처리 단계에서 데이터 형식 변환의 세부 과정은 다음과 같다.

① 데이터 형식 일치 정책 수립

같은 애트리뷰트명을 가지지만 데이터가 표현되는 방식이 일치하지 않는 경우, 하나의 통일된 기준을 정하여 데이터를 일관된 방식으로 표현한다. 예를 들어, 번호를 표현하는 방식에 있어 010-1111-2222 같이 '-'로 번호를 나누어 표현하는 경우와 01011112222와 같이 숫자로 꼭 나열하는 경우가 있다. 이러한 경우에는 둘 중에 하나로 번호를 표현하는 방식을 지정하여 애트리뷰트를 일관된 방식으로 표현한다.

② 자료를 구성하는 애트리뷰트의 개수가 다른 경우

기업의 업무 특성에 따라 하나의 자료를 저장하기 위한 애트리뷰트 수를 다르게 설계한다. 하나의 내용을 세부화하여 여러 애트리뷰트에 나눠서 저장하거나 반대로 여러 내용을 하나의 애트리뷰트에 합쳐서 저장하는 경우이다. 예를 들어, 주소를 표현할 때 필요에 의하여 '대주소', '상세주소', '우편번호' 등으로 나누어서 저장하는 경우와 하나의 애트리뷰트에 주소를 한꺼번에 저장하는 경우가 있을 수 있다. 이러한 자료의 처리 효율성을 판단하여 통일된 방법으로 애트리뷰트 수를 동일하게 조정한다.

③ 불필요한 자료 삭제

데이터를 일관된 방식으로 표현하기 위하여 필요한 내용만 간추리거나 단순화하는 작업이 필요하다. 이를 위해 저장된 데이터 중 불필요한 내용을 삭제하거나 수정한다. 예를 들어, 상품의 옵션을 나타내는 '★한정특가★고구마 샐러드1kg'에서 '★한정특가★'와 같은 불필요한 수식어를 제거한 후 '고구마'라고 표현할 수 있다.

4. 구현결과

4.1 구현 환경

이 논문에서 제안하는 고객 물류 데이터 통합 시스템을 구축하기 위한 환경은 아래 <표 1>과 같다.

<표 1> 시스템 구축 환경

H/W	프로세서	Intel i5
	메모리	8GB RAM
	네트워크	200MB/s
S/W	운영체제	Window 10
	라이브러리	xls-to-json, json-simple

4.2 통합 데이터베이스 스키마

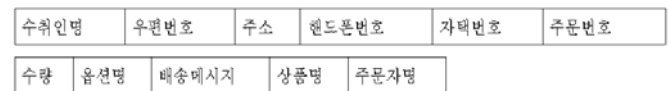
통합 데이터베이스 스키마는 여러 기업의 고객 물류 데이터의 애트리뷰트 중 같은 의미를 가지나 다른 애트리뷰트명을 가지는 애트리뷰트들을 기준을 정해 한 가지 애트리뷰트명으로 통일시켜 표현한다.

다음 <표 2>는 기업의 고객 물류 데이터에서 공통 애트리뷰트를 추출한 결과이다. 각 기업에서 “받는분 이름”, “받는 사람”, “수취인 이름”으로 표현한 애트리뷰트는 “수취인 명”이라는 공통 애트리뷰트로 정의하였다.

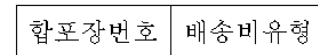
<표 2> 공통 애트리뷰트

공통	A기업	B기업
수취인명	받는분 이름	받는 사람
우편번호	우편번호	우편번호
주소	주소	주소
핸드폰번호	받는분 휴대폰1	전화번호1
자택번호	받는분 휴대폰2	전화번호2
주문번호	주문번호	주문번호
수량	구매수량	개수
옵션명	옵션명	옵션
배송메시지	배송메시지	배송메시지
상품명	상품명	상품명
주문자명	구매자이름	주문자명
X	합포장번호	업체상품코드
X	배송비유형	상품코드

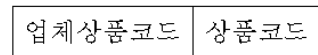
3개의 기업을 고객 물류 데이터를 분석하여 생성된 통합 데이터베이스의 스키마 구조는 (그림 4)와 같다. (가)는 공통 3개의 기업이 공통적인 요소를 찾아서 생성한 공통 스키마 이다. (나), (다)는 각 기업만이 가지고 있는 고유의 애트리뷰트로 구성된 스키마이다.



(가) 통합 데이터베이스의 공통 스키마



(나) 통합 데이터베이스의 A기업 고유스키마

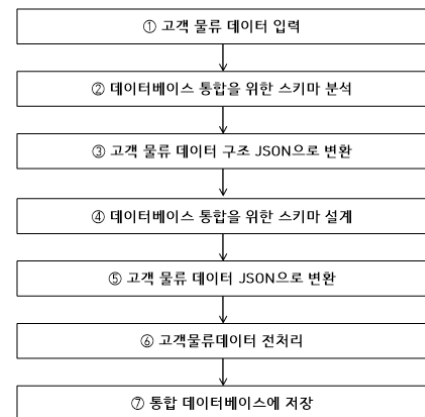


(다) 통합 데이터베이스의 B기업 고유스키마

(그림 4) 통합 데이터베이스 스키마 구조

4.3 시스템 화면

(그림 5)는 본 논문에서 제안하는 시스템에 대한 순서도 이다.



(그림 5) 전체 시스템 순서도

- ① 고객 물류 데이터를 입력받는다. 이때, 고객 물류 데이터는 데이터베이스가 아닌 Excel 형식의 파일이다.
- ② 데이터 스키마를 분석하여 서로 다른 고객 물류 데이터 구조의 공통 애트리뷰트들만을 모은다.
- ③ 입력으로 받은 고객 물류 데이터 구조를 (그림 2)와 같이 JSON 형태로 변환한다.
- ④ 통합 데이터베이스 스키마를 (그림 4)와 같이 설계한다.
- ⑤ 고객 물류 데이터를 (그림 3)과 같이 JSON 형태로 변환한다.
- ⑥ 데이터 형식을 일치시키고, 자료를 구성하는 애트리뷰트의 개수가 다른 경우 통일시키고, 불필요한 자료는 삭제한다.
- ⑦ 전처리 과정을 거친 JSON을 데이터베이스에 저장한다. 이때, 통합 데이터베이스 스키마에 포함되지 못하는 애트리뷰트들은 (그림 4)의 (가), (나)와 같이 고유의 데이터베이스 스키마에 따라 저장한다.

4.4 구현 결과

(그림 9)은 제안된 고객 물류 데이터 통합 시스템의 구현 결과이다.

The screenshot shows a web interface with a search bar and a table of results. The table has columns for '수출번호', '주소', '전화번호', '차량번호', '주요번호', '수량', '물류명', '배송회사', '발행일', and '주요사항'. There are three rows of data. Below the table, there is a section titled '기타 ASL 고유 데이터' with columns for '수출번호' and '주요번호', containing three rows of data.

(그림 9) 고객 물류 데이터 통합 시스템 구현 결과

고객 물류 데이터를 입력으로 받아 스키마 설계 모듈과 데이터 처리 모듈을 거쳐 생성된 통합 데이터베이스 결과를 (그림 8)과 같이 웹 페이지에 구현하였다.

결과 화면의 첫 번째 테이블은 통합 데이터베이스의 공통 스키마에 따른 데이터로 이루어져있고, 두 번째 테이블은 각 기업의 데이터베이스의 고유 스키마에 따른 데이터로 이루어져있다.

5. 결론

소규모의 기업에서는 비용적 부담으로 ERP 시스템을 도입하기에 한계가 있어 회사 운영에 필요한 작업을 Excel을 이용하여 수작업으로 처리하는 경우가 많다. 그로 인하여 업무를 처리하는 데 있어서 비효율성과 생산성 저하의 문제점이 존재하였다.

본 논문에서는 수작업으로 업무를 처리하는 데에서 오

는 비효율성과 생산성 저하의 문제를 해결하기 위해 다양한 형식의 데이터베이스를 하나로 통합하는 자동화 시스템을 구현하였다. 그 결과 수작업으로 진행하던 기존 작업 방식에서 Excel 파일만 입력하면 통합된 데이터베이스를 얻을 수 있는 작업 방식으로 개선되었다.

향후 연구에서는 통합 데이터베이스 모듈의 자동화 범위를 인터넷 전자상거래 사이트의 데이터베이스 뿐 아니라 다양한 사이트들의 데이터베이스로 확장하는 연구가 필요하다.

참고문헌

[1] Ramez Elmasri, Shamkant B. Navathe, “Database Systems Models, Languages, Design, and Application Programming”, 2014.

[2] 고영만, 배경제, “상이한 데이터 구조의 데이터베이스 간 통합 운영방안 연구 - 기초학문자료센터를 중심으로”, 한국문헌정보학회 추계학술대회, pp.3-4, 2011.

[3] D. Crockford, “The application/json Media Type for JavaScript Object Notation (JSON)”, 2006.

[4] 김기용, 김상대, 고소향, 류연승, 최승훈, 박우창, “데이터베이스 통합을 위한 전역 스키마 설계”, 한국멀티미디어 학회 학술발표논문집, pp.646-649, 2006.

[5] 윤선희, “복합인터넷쇼핑몰기반 데이터베이스의 설계”, 한국컴퓨터정보학회 학술발표논문집, Vol. 16, No. 2, pp.325-328, 2009.

[6] 김기중, “통합 데이터베이스를 위한 스키마 통합 방법”, 한국컴퓨터정보학회 학술발표논문집, pp.2-3, 1998.