

분산데이터베이스통합을위한에이전트시스템설계

강무홍^a, 최형림^b, 김현수^c, 홍순구^d, 최일용^e, 서영준^f

동아대학교 경영대학 경영정보과학부

부산 사하구 하단동 840, 604-714

^aTel: +82-51-200-7477, Fax: +82-51-207-2827, E-mail: mongy@dau.ac.kr

^bTel: +82-51-200-7477, Fax: +82-51-207-2827, E-mail: hrchoi@dau.ac.kr

^cTel: +82-51-200-7478, Fax: +82-51-200-7481, E-mail: hskim@dau.ac.kr

^dTel: +82-51-200-7488, Fax: +82-51-200-7481, E-mail: shong@dau.ac.kr

^eTel: +82-51-200-7488, Fax: +82-51-200-7481, E-mail: jandy@donga.ac.kr

^fTel: +82-51-200-7477, Fax: +82-51-207-2827, E-mail: chris@donga.ac.kr

Abstract

최근 새로운 수송망 설계를 통한 물류비 절감이 물류 관련 업체의 큰 이슈가 되고 있다. 하지만 일반적으로 행해지고 있는 수송망 설계는 대부분이 선박, 항공기, 열차 각각에 대해서만 고려되었으며 여러 가지 운송모드를 고려하는 복합운송망 설계는 극히 일부에 지나지 않았다. 이는 선박, 항공기, 열차를 모두 고려하여 수송망을 설계하기 위해서 필요한 운항 스케줄 정보가 산재되어 있어 복합운송망을 설계하기에는 많은 어려움이 따랐기 때문이다. 이러한 문제점을 해결하기 위해서는 모든 운송모드의 스케줄 데이터를 통합하여 제공할 수 있는 통합 데이터베이스 구축이 필요하다. 본 연구에서는 선박, 항공기, 열차에 대한 스케줄 정보를 제공하고 있는 정보원들을 분석하고 데이터 제공 방식을 조사하여 이를 모두 수용함으로써 데이터를 실시간으로 수집할 수 있는 에이전트 시스템의 설계를 수행하였다.

Keywords: 에이전트 시스템, 데이터 통합, 수송망 설계

1. 서론

컴퓨터와 인터넷의 보급에 따라 기존 오프라인에서 행해지고 있던 기업 내부 업무 및

상거래 업무의 온라인화가 급속히 증대되고 있다. 이는 업무 효율성 증대를 가져왔고 이에 더 많은 분야에서의 온라인화가 가속화되고 있는 추세이다. 이러한 업무의 온라인화는 다양한 분야에서 이루어지고 있고 특히 기업 내부 업무시스템 구축이 대표적인 예이다. 기업들의 기존 내부 업무시스템은 최초 메인프레임급 서버를 통해 사용되었던 터미널 방식의 프로그램을 시초로 개발되기 시작하였으며 클라이언트/서버 방식의 개발을 거쳐 현재는 시간과 장소에 구애를 받지 않고 업무를 처리할 수 있도록 인터넷 웹서비스 방식으로 많이 개발되고 있는 추세이다.

개발 방식의 변화뿐만이 아니라 적용 업무 분야에도 많은 변화가 있었다. 컴퓨터 기술은 최초 EDPS(Electronic Data Processing System)라 불렀던 것처럼 데이터의 처리를 위해서 주로 사용되었으나 현재는 기업 내의 거의 모든 업무 지원뿐만 아니라 기업 업무 통합, 신규 마켓 창출 등에 사용되고 있다.

이러한 기술의 발전은 3자물류 업계에도 영향을 미치고 있다[4]. 3자물류 업계에서는 주로 화주를 대신하여 화물을 출발지에서 도착지까지 전달해주는 업무를 맡고 있으며 업무 처리를 위해서 먼저 출발지, 도착지간의 운송 장비, 출발/도착 스케줄, 그리고 운임 등을 선정하는 최적 수송 네트워크를 설계하고 다음으로 선정된 운송 장비의 소유주(선사, 항공사

등)와 Booking 업무 등이 진행된다. Booking을 위해서는 많은 서류작업이 필요하며 이를 지원하기 위한 시스템으로는 문서관리시스템인 양재IT솔루션, WebLogis 등이 있다. 하지만 최적 수송 네트워크 설계를 지원하기 위한 시스템으로는 단순히 선박, 항공기, 철도 등의 스케줄 정보를 검색할 수 있는 Shipping Gazette, Schedule Bank 등 단순 검색시스템이 대부분이었으며 출발지, 목적지에 따른 최적 수송 네트워크를 찾아주는 시스템은 전무하였다. 현재 싱가포르의 전문 물류업체인 Accord사에서 이와 유사한 시스템을 개발하고 있다.

최적 수송 네트워크 선정이 어려운 이유는 크게 두 가지로 구분될 수 있다. 첫 번째는 수많은 수송 가능 물류망 중 최적을 선정할 수 있는 알고리즘 개발의 어려움을 들 수 있고 두 번째로는 네트워크를 선정하기 위한 스케줄 데이터 확보의 어려움을 들 수 있다. 국제 운송은 일반적으로 선박을 이용하여 운송되고 있으나 급한 운송의 경우는 항공기를 이용할 수도 있으며 때에 따라서는 선박, 항공기, 열차를 모두 이용하는 경우도 있다. 따라서 선박, 항공기, 열차의 모든 스케줄이 필요하며 이는 많은 노력과 데이터 수집기술을 필요로 한다.

현재 선사, 항공사, 철도공사는 다양한 방식으로 스케줄 정보를 제공하고 있다. 일반적으로 각 사에서 운영하고 있는 웹사이트를 통해서 제공하고 있으나 이는 실시간으로 데이터

를 연계할 수 없다. 하지만 이들은 웹사이트뿐만 아니라 엑셀 파일 또는 XML/EDI 형태로도 정보를 제공하고 있다.

위의 <Table 1>은 각 업체의 스케줄 담당자와 직접 인터뷰를 수행하거나 인터넷 웹사이트를 분석하여 조사하였다.

본 연구에서는 최적 수송 네트워크 선정을 지원하기 위해 필요한 각종 스케줄 데이터들을 수집하기 위해 에이전트 기술을 적용하였다. 이는 Table-1과 같이 다양한 환경을 가지고 있는 정보원들을 고려하여 다양한 방식의 스케줄 데이터를 수집하기 위함이며 이를 위한 수집 기술을 에이전트에 접목하였다.

2. 관련 시스템 현황

인터넷 상에는 이미 선박, 항공기, 열차 스케줄을 제공하고 있는 웹사이트들이 많이 존재하고 있다. 선박 스케줄 정보를 제공하고 있는 Shipping Gazette, 항공기 스케줄 정보를 제공하고 있는 Traxon Korea가 대표적이며 이들은 단일 운송 장비의 스케줄만을 제공하고 있다. 따라서 선박, 항공기를 고려한 복합 운송 네트워크를 설계하기 위해서는 두 사이트에 모두 접속하여 스케줄을 검색하여야 하고 스케줄 연결을 사용자가 직접해야 한다는 한계점이 발생된다.

30년간 축적된 선박 스케줄 데이터를 통해 국내 최대의 선박 스케줄 정보원으로 자리 잡고 있는 Shipping Gazette는 선박 스케줄뿐만 아니라 항만·물류와 관련된 다양한 정보를 제공하고 있다. 이러한 스케줄 데이터 수집을 위해 Shipping Gazette는 각 선사로부터 일반 문서 또는 엑셀 파일을 팩스, 이메일과 같은 매체로 수집하여 직접 입력하고 있다. 따라서 불필요한 입력 작업이 필요하며 입력에 따른 시간적, 비용적 낭비가 발생한다.

Traxon Korea는 대한항공을 포함한 전 세계의 16개 항공사가 공통 주주로 설립한 회사로써 36개의 항공사 운항 스케줄을 제공하고 있다. 운항 스케줄 외에도 항공화물에 대한 Booking과 Tracking 서비스를 제공하고 있으

Table 1 - 각 정보원들의 스케줄 제공 방식

업체분류	업체명	제공 방식			
		웹	엑셀	EDI	DB
선사	한진해운	○	○		○
	P&O Nedlloyd	○		○	
	현대상선	○	○		
	태영상선	○			
항공사	진양해운	○			
	대한항공	○	○		
철도공사	아시아나항공	○	○		
	한국철도공사	○	○		○
기타 정보원	Shipping Gazette	○			
	Schedule Bank	○	○		
	Traxon Korea	○		○	



Figure 1 - Shipping Gazette

며 현재 XML 웹서비스를 통한 스케줄 정보 제공, Booking/Tracking 연계 서비스를 제공하기 위해 기술 개발 중에 있다. Traxon Korea는 각 항공사의 데이터를 통합 Repository에 저장하지 않고 필요한 데이터 요청 및 수집을 UN/EDIFACT과 CARGO-IMP의 혼합형 EDI 문서를 통해 수행하고 있다. 따라서 잦은 데이터 요청에 따른 과부하가 야기될 수 있으며 전송 및 수신에 대한 시간 낭비¹⁾가 초래된다.

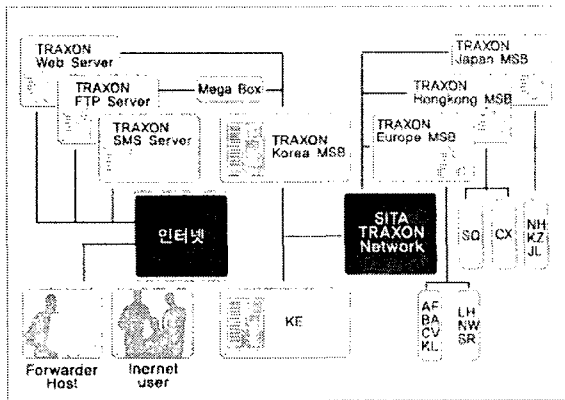


Figure 2 - Traxon Korea 시스템 구조

3. 기존 데이터 통합 기술 현황

본 장에서는 데이터 연계를 위한 기존 기술에 대해서 분석해 보고 이들의 문제점을 파악하여 이를 해결할 수 있는 방안을 제시한다.

데이터 연계 기술은 기업의 업무 통합을 위해 제시되었던 EAI(Enterprise Application

Integration) 환경을 기반으로 하는 미들웨어 기술이 있다. 본 기술은 미들웨어를 통해 여러 대의 데이터베이스 서버를 하나인 것처럼 통합하는 기술로써 데이터의 동기화를 목적으로 하고 있다. IBM, BEA 등과 같은 대형 벤더업체에서 본 기술을 개발/판매하고 있으며 고가, 고사양의 서버를 필요로 하여 중·소기업에서는 사용을 못하고 있다.

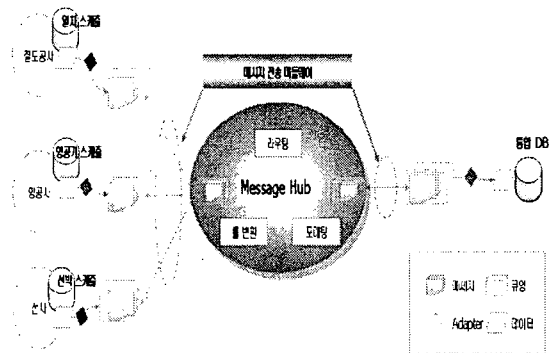


Figure 3 - EAI 기반 데이터 통합 구조

다음으로 Replication과 DTS(Data Transformation Service)가 있다. 이 기술들은 DBMS(Database Management System)에서 기본적으로 제공하고 있으며 주기적으로 데이터를 복제하거나 전송하는 기술이다. 하지만 본 기술은 데이터 Source가 있는 정보원 서버와 정보 수집을 하려하는 서버가 Trust가 되어 있거나 Source 정보원 데이터베이스에 대한 접근 권한이 있어야 데이터를 수집할 수 있다는 한계점이 있다.

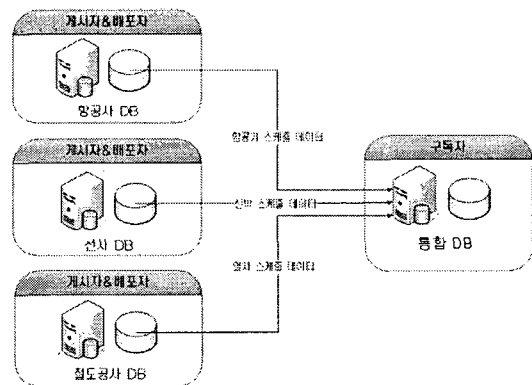


Figure 4 - Replication 데이터 수집 구조

1) Traxon Korea에서 각 항공사의 스케줄 데이터를 로딩하기 위해서는 5~7초정도 소요됨, Traxon Korea 인터뷰

본 연구에서는 이러한 기존 데이터 연계 기술들의 고비용/고사양 문제와 접근 권한의 문제를 해결하고 데이터를 수집할 수 있는 에이전트 시스템을 제안한다.

4. 데이터 수집 방안

1장에서 언급하였다시피 선박, 항공기, 열차 스케줄 정보를 각 정보원들은 웹사이트, 엑셀 파일, XML/EDI, 그리고 데이터베이스 연계를 통해서 제공하고 있음을 알 수 있었다. 본 연구진은 이렇게 다양한 데이터 제공방식을 모두 수용할 수 있는 에이전트를 개발하기 위해 연계 방안을 모색해 보았다.

Table 2 - 데이터 수집 방안

DB 접근성 가능여부	데이터 제공 방법	데이터 확보 기술
가능	동일 DBMS의 DB 복사	Replication(복제), 미들웨어
	이 기종 DBMS의 DB 전체 또는 일부 이관	Migration, 미들웨어
	이 기종 DBMS의 스케줄 데이터 수집	DTS, 미들웨어
불가능	문서 전송	직접 입력
	엑셀 파일 및 XML/EDI 전송	Excel 파일 및 EDI Parser

<Table 2>에서 보는 것처럼 본 연구에서는 데이터 수집 방안의 기준을 데이터베이스 접근 가능의 여부에 두었다. 정보원에서 데이터베이스 접근 권한을 제공하면 Replication/DTS 및 미들웨어와 같은 기존의 데이터 연계 기술을 적용하고 접근권한이 없을 경우에는 별도의 기술을 개발하여 이를 통해 데이터를 수집하는 방안을 수립하였다.

먼저 엑셀 파일을 수집하기 위해서는 파일 내부의 셀에서 데이터를 추출할 수 있도록 Jakarta Project의 POI 컴포넌트[10]를 이용하여 기술을 개발하였다.

<Figure 5>에서 보는 것과 같이 파일 업로드를 위한 인터페이스를 통해 엑셀 파일을 업로드하고 POI의 HSSF를 통해 각 셀의 데이

터를 추출하게 되며 이 데이터는 통합 Repository에 쿼리문으로 전송되게 된다.

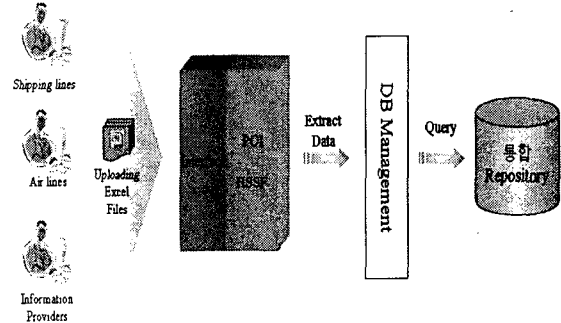


Figure 5 - 엑셀 Parser 구조

정보원에서 XML/EDI 형태로 데이터를 제공할 경우에는 XML 메시지를 Parsing할 수 있도록 DOM(Document Object Model) 또는 SAX(Simple API for XML)를 이용하여 XML Parser를 구현하였다. 이를 위해서는 정보원의 XML 구조를 분석하기 위한 DTD(Document Type Definition)를 먼저 분석하였으며 DTD 구조에 따라 자동으로 데이터를 추출하였다.

마지막으로 정보원의 DB에 직접 접근할 수 있는 경우에는 스케줄 데이터에 대해서 직접 쿼리할 수 있으므로 신규 데이터의 등록, 데이터의 수정을 감지하여 직접 데이터를 실시간으로 수집할 수 있는 기술을 개발하였다.

본 연구에서는 실시간 데이터 수집 에이전트 시스템을 개발하기 위해 앞서 언급한 엑셀 및 XML/EDI Parser, DB 직접 접근을 위한 기술들을 개발하고 1장의 <Table 1>의 정보원들에 적용해 보았다.

Table 3 - 기술 적용 대상 정보원

제공방식	Test-bed
엑셀	한진해운, 현대상선, 대한항공, 한국철도공사
XML/EDI	P&O Nedlloyd
DB 접근	Dummy DB 생성 후 이를 시험

기술 적용 결과로 엑셀 Parsing 기술의 경우 각 정보원마다 사용하는 엑셀 템플릿이 다

르기 때문에 정보원별로 Parser의 커스터마이징이 필요함을 알게 되었으며 이와는 달리 XML/EDI의 경우 DTD를 통해 XML의 구조를 파악할 수 있어 자동으로 데이터 추출이 가능하였다. 또한 DB 접근의 경우는 현재 Test-bed로 접근 권한을 줄 수 있는 정보원이 없어 별도의 Dummy DB를 구축한 후 이를 통해 기술의 적용을 수행하여 데이터의 수집이 실시간으로 가능함을 알 수가 있었다.

5. 에이전트 시스템 설계

4장에서는 다양한 환경을 가지고 있는 정보원들의 정보시스템에서 스케줄 데이터를 수집할 수 있도록 여러 가지 방안을 수립하고 세부 기술들을 개발/적용해 보았다. 본 장에서는 이러한 기술을 통해 정보원의 스케줄 데이터의 업데이트를 실시간으로 감지하고 스케줄 정보를 수집할 수 있는 에이전트 시스템의 설계를 수행하였다. 기 개발된 기술을 통합하는 것만으로도 데이터 수집이 가능하지만 파일 업로드, DTD 분석 등 사용자의 행동이 필요하기 때문에 자율적이고 능동적인 데이터 수집이 불가능하였다. 따라서 개발된 기술에 에이전트 기술을 접목하여 능동적인 데이터 수집을 위한 에이전트 시스템의 설계를 본 장에서 수행하였다.

데이터를 수집하기 위한 에이전트 시스템은 4장의 데이터 수집 기술 외에도 통신 기술, 파일 업데이트 감지 기술, 그리고 XML 메시지 구조 설계 및 Parsing 기술이 필요하다.

에이전트 통신을 위해서는 다양한 기술들이 등장하고 있다. 이 중 MAS (Multi-Agent System)는 에이전트간 협업을 지원하는 시스템으로 FIPA-OS, JADE(Java Agent DEvelopment Framework), MAFNS (Multi-Agent Framework for Negotiation System) 등의 Framework을 통해 용이하게 구현할 수 있다[7]. 본 연구에서는 MAFNS[7]를 이용하여 에이전트 통신환경을 구현할 수 있도록 시스템 구조를 설계하였다. 기 개발된 데이터 수집 기술과 파일 업데이트 감지 기술,

XML Parser는 MAFNS의 Agent 구조에서 Application Layer에 포함되어 데이터 수집을 수행하게 된다.

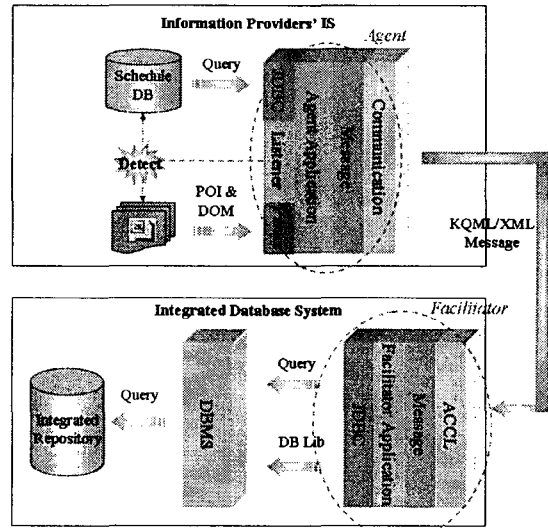


Figure 6 - 데이터 수집 에이전트 시스템 구조

<Figure 6>에서 보는 것과 같이 데이터 수집 에이전트는 Listener를 통해서 데이터베이스 또는 엑셀 파일, XML/EDI 파일의 등록을 감지하고 신규 데이터가 업데이트 되었을 경우 JDBC를 통해서 데이터베이스에서 데이터를 수집하거나 POI를 이용하여 엑셀 파일의 데이터를 추출하게 된다. XML/EDI의 경우도 XML 파일이 저장 폴더에 업데이트 되었음을 감지하고 XML Parser를 통해서 데이터를 추출하게 된다. 이렇게 추출된 데이터들은 스케줄 정보를 담은 XML 문서로 변경되게 되며 이는 에이전트 통신을 위한 메시지 통신 언어인 KQML(Knowledge Query and Manipulation Language)[6]의 Content에 포함되어 전송되게 된다. 스케줄 데이터 전송을 위한 XML 스키마는 아래와 같은 DTD로 표현될 수 있다.

<Figure 7>의 DTD에서 알 수 있듯이 XML 메시지는 크게 정보원의 기본 정보를 나타내는 information_provider 부분과 스케줄 데이터를 담고 있는 schedule 부분으로 나뉜다. 아래 <Figure 8>은 에이전트가 Facilitator로 전송하게 되는 KQML/XML 메

시지의 한 예이다.

```
<?xml version="1.0" encoding="euc-kr" ?>
<!ELEMENT schedule_message (information_provider, schedule*)>
<!ELEMENT information_provider (id, name)>
<!ELEMENT id (#PCDATA)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT schedule (schedule_type, vehicle_name, vehicle_id,
departure_location, departure_time, arrival_location, arrival_time,
pass*)>
<!ELEMENT schedule_type (#PCDATA)>
<!ELEMENT vehicle_name (#PCDATA)>
<!ELEMENT vehicle_id (#PCDATA)>
<!ELEMENT departure_location (#PCDATA)>
<!ELEMENT departure_time (#PCDATA)>
<!ELEMENT arrival_location (#PCDATA)>
<!ELEMENT arrival_time (#PCDATA)>
<!ELEMENT pass (pass_location, pass_arrival_time,
pass_departure_time)>
<!ELEMENT pass_location (#PCDATA)>
<!ELEMENT pass_arrival_time (#PCDATA)>
<!ELEMENT pass_departure_time (#PCDATA)>
```

Figure 7 - 스케줄 메시지의 DTD

```
(response
  :sender agent_hanjin
  :receiver facilitator
  :content
    (
      XML Message
    )
  :in-reply-to 0
  :language tnps_lang
  :ontology tnps
  :protocol fipa-contract-net
)
```

Figure 8 - KQML/XML 메시지 구조

에이전트 시스템에서는 KQML을 통해서 에이전트와 Facilitator 간의 메시지 통신을 수행하게 되며 실제 전송되어야 하는 스케줄 정보는 <Figure 7>의 DTD 구조를 가지고 있는 XML 메시지를 통해서 전달된다.

Facilitator는 전송받은 메시지에서 XML 부분을 Parsing하게 되고 쿼리문을 작성하여 통합 Repository에 저장하게 된다.

6. 결론

본 연구에서는 분산되어 있는 데이터들의 수집을 위해서 다양한 기술을 개발하고 이를 모두 포함함으로써 능동적으로 데이터를 통합할 수 있는 에이전트 시스템의 설계를 수행하였다. 본 연구를 기반으로 연구진은 실시간 데

이터 통합이 가능한 에이전트 시스템의 개발을 진행 중에 있다. 설계된 에이전트 시스템의 개발에 따라 저사양, 저비용으로 분산 데이터 베이스의 통합이 가능할 것이며 이를 통해 효율적인 데이터 관리가 가능할 것이다.

일반적으로 기업들은 데이터 관리를 DBMS를 통하여 수행하고 있고 간단한 데이터는 엑셀에 의존을 하고 있다. 하지만 워드프로세서 파일, 텍스트 파일 등의 파일 시스템으로 관리하거나 기타 다른 방식을 통해 데이터를 관리하고 있는 기업들도 존재하고 있다. 완전한 기업 데이터의 통합을 위해서 본 연구의 수집 기술 외의 데이터 수집 기술 개발에도 연구가 필요할 것이다.

References

- [1] Andrea, C., Diego, C., Giuseppe, D.G., Maurizio, L. (2004). "Data integration under integrity constraints," *Journal of Information Systems*, Vol.29, pp. 147 - 163.
- [2] Theodore, P.S., Patricia, J.D. (1997). "The impact of operating environment on the formation of cooperative logistics relationships," *Transportation Research Part E: Logistics and Transportation Review*, Vol. 33, No. 1, pp. 53-65.
- [3] Hertz, S., Alfredsson, M. (2003). "Strategic development of third party logistics providers," *Journal of Industrial Marketing Management*, Vol.32, No.2, pp. 139 - 149.
- [4] John, L.C., Gary, R.A., Mark, J.C. (2003) "Third-Party Logistics Study Results and Findings of the 2003 Eighth Annual Study" Technical Report, Georgia Institute of Technology.
- [5] Choi, H.R., Kim, H.S., Hong, S.G., Park, Y.J., Park, Y.S., Kang, M.H. (2005) "Implementation of Framework for

Developing Multi-Agent based Automated Negotiation Systems," *The 7th International Conference on Electronic Commerce*, pp. 306-315.

- [6] Tim, F., Yannis, L. James, M. (1997) "KQML as an agent communication language," Technical Report, F49620-92-J-0174, Dept. of Computer Science, University of Maryland Baltimore Country.
- [7] 최형립, 김현수, 박남규, 박영재, 김성훈, 이현철 (1999) "수출입 컨테이너화물 통합 데이터베이스 구축", 한국전자거래학회, 한국정보시스템학회 종합학술대회 논문집.
- [8] 정지호, 윤청 (2004) "EAI와 Web Service 환경에서 트랜잭션의 효율적인 처리 방안", 정보처리학회지, 제11-D권, 제2호.
- [9] 박진수, 김성현 (2004) "Real Time Enterprise 구현을 위한 데이터 통합 방안 연구", 한국경영정보학회, 2004 추계학술대회, 172-178.
- [10] Jakarta POI, <http://jakarta.apache.org/poi>