The Journal of Digital Policy & Management 2013 Oct; 11(10): 47-57

# 빅데이터 패키지 선정 방법

변대호
경성대학교 경제금융물류학부

# Method for Selecting a Big Data Package

**Dae-Ho Byun**
**Dept. of Economics, Finance, and Logistics**

**요 약** 빅데이터 분석은 데이터의 양, 처리속도, 다양성 측면에서 데이터 마이닝과 달리 문제해결과 의사결정을 위해서는 새로운 도구를 필요로 한다. 많은 글로벌 IT기업들은 사용하기 쉽고 기능성이 우수한 모델링 능력을 가진 다양한 빅데이터 제품을 출시하고 있다. 빅데이터 패키지는 분석도구, 인프라, 플랫폼 형태로 하드웨어와 소프트웨어를 포함한 솔루션이다. 빅데이터의 수집, 저장, 분석, 시각화가 가능한 제품이다. 빅데이터 패키지는 업체별로 제품 종류가 많고 복잡한 기능을 가질 뿐만 아니라 선정에 있어서 전문 지식을 필요로 하며 일반적인 소프트웨어 패키지보다 그 중요성이 높기 때문에 의사결정 방법의 개발이 요구된다. 본 연구는 빅데이터 패키지 도입을 위한 의사결정지원 방법을 제안하는 것이 목표이다. 문헌적 고찰을 통하여 빅데이터 패키지의 특징과 기능을 비교하고, 선정기준을 제안한다. 패키지 도입 타당성을 평가하기 위하여 비용과 혜택 각각을 목표노드로 하는 AHP 모델 및 선정기준을 목표노드로 하는 AHP 모델을 제안하고 이들을 결합하여 최적의 패키지를 선정하는 과정을 보인다.

**주제어** : 빅데이터 분석, 패키지, 솔루션, 선정기준, AHP, 의사결정

**Abstract** Big data analysis needs a new tool for decision making in view of data volume, speed, and variety. Many global IT enterprises are announcing a variety of Big data products with easy to use, best functionality, and modeling capability. Big data packages are defined as a solution represented by analytic tools, infrastructures, platforms including hardware and software. They can acquire, store, analyze, and visualize Big data. There are many types of products with various and complex functionalities. Because of inherent characteristics of Big data, selecting a best Big data package requires expertise and an appropriate decision making method, comparing the selection problem of other software packages. The objective of this paper is to suggest a decision making method for selecting a Big data package. We compare their characteristics and functionalities through literature reviews and suggest selection criteria. In order to evaluate the feasibility of adopting packages, we develop two Analytic Hierarchy Process(AHP) models where the goal node of a model consists of costs and benefits and the other consists of selection criteria. We show a numerical example how the best package is evaluated by combining the two models.

**Key Words :** Big data analysis, package, solution, selection criteria, Analytic Hierarchy Process, decision making

* 이 논문은 2013학년도 경성대학교 학술연구비지원에 의하여 연구되었음.
Received 1 August 2013, Revised 22 August 2013
Accepted 20 October 2013
Corresponding Author: Dae-Ho Byun(School of Economics, Finance, and Logistics)
Email: dhbyun@ks.ac.kr

ISSN: 1738-1916

## 1. 서론

2000년대 이후 IT환경의 혁신은 의사결정 문제를 해결하기 위해서 모델링이나 수학적 접근 보다는 데이터를 이용하는 방법이 요구되고 있다. 저장장치나 CPU 기술의 비약적 발전은 과거에는 불가능하였던 고용량, 빠른 속도, 지능화되어 테라바이트 이상의 빅데이터 분석이 가능하게 되었다. 빅데이터 분석은 비즈니스뿐만 아니라 공공분야에서 니즈가 증가되고 있다. 시장조사기관인 IDC에 따르면 세계 빅데이터 시장은 매년 39.4%씩 성장해 2015년 169억달러 규모로 증가할 것이라고 전망했다. 위키본(Wikibon)은 빅데이터 시장 규모가 2012년 51억달러에서 2017년 534억달러로 보다 높은 성장률(연평균 60%)에 이를 것으로 예상했다. 전자신문[8]은 한국과학기술정보연구원 조사에 따라 국내 빅데이터 시장규모는 오는 2015년 2억6300만달러, 2020년께는 8억500만달러(약 9000억원)에 이를 것으로 예상했다.

빅데이터 분석은 예측기법, 데이터마이닝, 통계학, 인공지능, 자연어 처리에 기반을 두고 고용량의 데이터를 빠른 속도로 분석할 수 있는 도구들의 집합으로 빅데이터의 수집, 저장, 처리, 분석, 시각화 과정을 거쳐 의사결정 결과를 도출하는 전 과정을 의미한다. Brown, Chui, Manyika[3], Brynjolffson, Hammerbacher, Stevens[4]는, 빅데이터 분석이 중요한 이유는 데이터로부터 유용한 지식을 도출하는 일은 비즈니스에서 끊임없는 도전이며 기업의 성과와 경쟁적 우위를 달성할 수 원천이기 때문이라고 했다. Chen, Chiang, Storey[5]은 빅데이터 분석은 모바일 기기와 각종 센서로부터 발생하는 콘텐츠를 처리하는 비즈니스 인텔리전스와 분석의 3.0 단계로 규정하고 데이터, 텍스트, 웹, 네트워크, 모바일 분석을 포함한다고 했다. 과거 데이터웨어하우스, 데이터마이닝 기법을 이용한 비즈니스 인텔리전스 프로그램과 공통점이 있지만 빅데이터 분석이 가진 차이점은 첫째, 전자는 여러 개의 상세함을 살펴보는 것이 목적이라면 후자는 데이터의 관계성에 초점을 둔 큰 그림을 도출하는 것이다. 큰 그림에서 영향력을 행사하는 엔티디를 밝혀내는 것이 관심사이다. 둘째, 전자보다 후자가 처리해야할 데이터 크기가 테라바이트에서 페타바이트까지 엄청나게 크다는 점과 처리절차가 복잡하며 처리해야할 데이터 유형이 다양하다. 숫자나 문자 형태의 파일뿐만 아니라 소셜미디어나 웹사이트 내의 텍스트, 이미지, 비디오, RFID에서 발생하는 트랜잭션 데이터 까지를 포함하게 된다. 셋째, 데이터마이닝 분석은 상세성에 목적을 두기 때문에 모델의 종류가 많을 뿐만 아니라 모델링에 시간이 걸리지만 빅데이터 처리는 처리언어를 사용하여 자동화를 목표로 한다는 점이다.

IT 환경적 변화로 하둡(Hadoop), 맵리듀스(MapReduce), R 프로그래밍, NoSQL의 개발은 빅데이터 분석을 가능하게 하였다. White[31], Stonebraker, et al.[28]은 하둡은 빅데이터 분석에 보편적으로 사용되는 대용량의 데이터 처리를 위해 개발된 오픈소스 소프트웨어이며, 맵리듀스는 분산데이터베이스 처리를 위한 마이닝 알고리즘으로 정의하였다. NoSQL은 과거 SQL을 표준으로 하는 관계형 데이터베이스와 달리 테이블간 조인이 필요 없고, 스키마가 유동적이어서 빅데이터 처리에 유용하다. 또한 대용량의 데이터를 저장할 수 있으며 데이터를 여러 대의 서버에 분산 저장하기 때문에 서비스의 안정성이 높다. Venables, Smith, R Core Team[29]는 R프로그래밍 언어는 오픈소스로 통계기법, 수치해석, 시각화 기능을 가지며 사용하기 쉽다는 점 때문에 빅데이터 분석에 유용하다고 했다.

Russom[24]에 따르면 빅데이터 분석을 위해서는 많은 기법과 프로그램이 있기 때문에 프로젝트 성격에 따라 벤더가 제공하는 도구 유형, 사용자 기법과 방법 등을 적절히 선택해야 하며, 여기에는 소프트웨어와 하드웨어의 평가를 포함한다. Russom[24]에는 실증조사를 통해 맵리듀스, 하둡, 텍스트마이닝, 클라우드, No-SQL 등 33개의 빅데이터 분석 기법과 도구의 중요도를 도출하였다. 이러한 빅데이터 분석 기술은 하드웨어와 소프트웨어를 포함한 패키지 형태로 구현되며 이는 분석함수, 플랫폼, 인프라, 도구 등을 내포한 솔루션을 의미한다. EMC, IBM, HP, 오라클, LG 등 국내외적으로 많은 IT기업들이 빅데이터 패키지를 출시하고 있으며 최근에는 전문 중소기업들까지 동참하고 있다. 구글은 빅쿼리 서비스라는 저장공간과 분석 솔루션을 제공한다. 하드웨어 업체로는 HP, 델 등이 빅데이터 분석용 서버를 출시하도 있다. LG CNS는 하드웨어 소프트웨어 일체형인 스마트 빅데이터 플랫폼 어프라이언스, 삼성SDS의 빅데이터 워크플로우 자동화 솔루션인 애널리틱스 파운데이션을 출시했다.

빅데이터 분석을 위해서는 특별한 프로그램이 필요한

데 본 연구에서는 이를 빅데이터 패키지 로 정의한다. 빅데이터 패키지는 상업용 목적으로 특정 업체들이 개발한 분석도구, 플랫폼, 소프트웨어, 하드웨어 인프라를 포함한 통합 제품으로 정의한다. 업체별로 출시되는 제품이 다양하고 복잡한 기능을 가지므로 패키지 선정은 전문적 지식을 필요로 한다. 또한 데이터베이스, ERP, CRM 솔루션, 데이터 마이닝 도구 등 일반적인 패키지 선정문제와는 달리 그 중요성이 높다. 첫째, 빅데이터 패키지는 소프트웨어뿐만 아니라 하드웨어 선정까지를 포함하고 데이터 수집에서 시각화에 이르는 전 과정을 지원할 수 있어야 한다. 둘째, 빅데이터는 조직에서 전략적인 의사결정을 내리는 문제에 사용될 가능성이 높다. 셋째, 내부 데이터를 활용할 것인지 웹이나 소셜 데이터를 활용할 것인지에 따라 도입 목적은 다르기 때문에 활용 목적과 범위에 따라 가장 쉽고 효과적이며, 효율적으로 달성할 수 있어야 한다. 넷째, 사용에 전문성을 요구하기 때문에 교육훈련을 통해 분석가를 양성하고 장기 지속적으로 사용되어야 한다. 다섯째, 빅데이터 분석의 중요성으로 인하여 패키지 선정 역시 중요하다. 최적의 패키지 선정은 국가경제와 기업성과에 영향을 미치므로 비용과 기대효과를 고려한 도입 타당성에 대한 분석이 요구된다. 이투데이[7]에 따르면 맥킨지는 빅데이터 분석은 미국 GDP 성장을 이끌 수 있다고 했다. 연간 리테일과 제조업 GDP에 기여하는 규모는 3250억 달러에 달할 수 있다. 또 보건과 정부서비스 비용을 최대 2850억 달러 줄일 수 있다고 했다. 여섯째, 빅데이터 기술은 상당부분 오픈소스에 기반을 두고 발전되고 있기 때문에 확장성을 고려해야 한다.

Jadhav, Sonar[15]에 다르면 소프트웨어 패키지 선정 문제는 선정절차, 선정기준, 선정방법을 해결하는 것이라고 했다. 지금까지 의사결정 문제를 해결하기 위하여 많은 패키지 선정 기준과 방법들이 논의되었다. Gürbüz, Alptekin, Alptekin[9]은 ERP 패키지, Alshawi, Missi, Irani[1]는 CRM 솔루션, Seng, Chen[26]은 데이터마이닝 솔루션 선정 방법을 기술하였다. 빅데이터 분석용 패키지 선정 문제는 새로운 선정기준과 선정절차를 제시해야 한다는 면에서 이들 연구와는 차별화 된다. 선행연구들은 패키지 선정방법으로 선정기준의 가중치를 반영하여 최적의 대안을 선정하는 다기준의사결정 방법[22] 들을 사용하였다.

본 연구는 빅데이터 패키지 선정 방법을 제안하는 것이 목적이다. 먼저 선정 대안인 빅데이터 분석용 도구, 솔루션, 인프라 현황과 패키지 형태로 출신된 제품의 기능을 고찰한다. 선정기준은 일반적인 소프트웨어 패키지들의 선정기준, 의사결정지원시스템(DSS) 소프트웨어, 데이터마이닝 도구의 선정기준들, 상업용 빅데이터 패키지들이 가진 특징으로부터 도출하기로 한다. DSS는 모델을 구축하여 솔루션을 찾는 다는 점, 데이터마이닝 도구는 유용한 지식을 추출할 수 있는 알고리즘이 내장되어 있다는 점에서 빅데이터 분석과 유사하다. 그리고 상업용 빅데이터 패키지들은 사용자의 요구사항이 반영되어 있다고 볼 수 있기 때문이다.

그리고 실제 빅데이터 패키지를 도입하려는 기업은 여러 업체의 솔루션 비교하여 장점을 가진 업체 솔루션들을 선별적으로 선택하여 사용하고자 할 것이다. 그러나 본 연구에서는 한 업체에서 출시한 여러 가지 솔루션을 하나의 패키지로 간주하여 모든 솔루션을 선정하는 문제로 전제한다. 기업에서 여러 업체의 솔루션을 사용하는 것은 기존 시스템과의 호환성 문제, 유지보수, 교육훈련 등 복잡한 문제가 있게 되고 빅데이터 프로젝트가 대부분 초기 단계이므로 특정 업체의 솔루션을 먼저 사용해보는 것이 일반적인 접근이기 때문이다.

선정기준으로부터 최적의 대안을 선정하기 위하여 Saaty, Kearns[25]이 제안한 계층적분석과정(AHP) 방법을 사용한다. AHP 모델은 2개로 구축된다. 도입타당성을 판단하기 위해 비용과 혜택을 목표노드로 갖는 AHP 모델로부터 후보 대안을 선정한 후 도출된 선정기준을 목표노드로 갖는 AHP 모델과 최적의 패키지를 선정하는 모델을 제안한다. 빅데이터 패키지 도입이 조직에 미치는 영향이 크기 때문에 비용과 혜택을 선정기준과 같은 수준으로 두고 세부기준에 대한 가중치를 조정하는 방법보다는 2개의 AHP 모델 구축이 바람직하다. 끝으로 본 연구에서 제안된 모델을 수치적인 예를 통해 구현한다.

## 2. 빅데이터 패키지 현황

### 2.1 벤더

빅데이터 솔루션은 기존 하드웨어나 소프트웨어 개발회사들이 주축이 되어 개발하거나 데이터마이닝 업체, 신규 기업들이 진출하고 있다. 데이터마이닝 도구의 발

전적 형태 또는 하둡이나 R과 같은 오픈소스를 활용하는 경우, 그리고 각종 프로그램을 이용한 플랫폼, 하드웨어 서버 등 매우 다양한 형태를 보이고 있다. 솔루션별로 기능이 다양하여 하나의 패키지가 모든 의사결정 문제를 해결하기는 어려운 점이 있다. 빅데이터 분석의 전제조건은 데이터 수집이 가능해야 하고 활용할 데이터가 존재해야 한다는 점 때문에 수집 가능한 데이터와 분석 가능한 솔루션이 존재해야하며 패키지는 해결하려는 문제 유형을 지원할 수 있어야 한다.

정지선[16]에 따르면 빅데이터 사업을 추진 중인 글로벌 IT기업인 EMC, HP, IBM, 오라클, SAS, 테라데이타에 대하여 상업용 솔루션을 조사하였다. 대표적 솔루션으로 EMC의 저장 솔루션, 콘텐츠 관리 솔루션과 HP는 대량데이터, 실시간 데이터 분석이 가능한 버티카 플랫폼, 의미 분석이 가능한 오토노미 솔루션을 들 수 있다. 그리고 IBM의 인포스피어, 빅인사이트, 오라클의 빅데이터 어플라이언스, SAS의 솔루션맵, 테라데이타의 애스터맵리듀스 플랫폼을 들 수 있다.

또한 국내 IT기업으로는 그루터, 넥스알, 다음소프트, 사이람, 솔트룩스 등이 제품을 출시하고 있다. 그루터는 빅데이터 플랫폼인 BAAS를 통해 데이터의 수집, 저장, 분석을 위한 컨설팅 서비스와Seenal.com은 소셜네트워크 데이터 수집 및 분석서비스를 제공한다. 넥스알의 NDAP 플랫폼은 빅데이터 분석 및 실시간 질의가 가능하며, 데이터웨어하우스 기반인 오픈소스인 RHive 플랫폼으로는 고급 분석이 가능하다. 다음소프트는 소셜미디어 분석을 위해 SOCIALmetrics, TrendMap 솔루션을 출시했다. 사이람은 소셜 네트워크 분석용 소프트웨어인 NetMiner는 백만개의 노드와 천만개의 링크를 분석하고 시각화가 가능한 제품이다. 솔드룩스는 비정형 데이터 분석이 가능한 트루스토리, 시맨틱 검색용 IN2, 추론용 STORM 및 분석 서비스 플랫폼인 O2를 제공한다.

김수지, 이재희[17]은 국내 빅데이터 솔루션 기업으로 센솔로지, 아크원소프트, 알테어, 여인소프트 등 14 업체를 조사하였다. 국내기업들의 특징은 중소 벤처기업 위지로 가격 경쟁력을 가진 국산 솔루션을 출시하고 있으며 관련 기관 및 기업간 협업체계를 구축하고 있다. 그리고 각 기업별로 특화된 솔루션을 제공한다는 것이 특징이다. 센솔로지의 솔루션으로는 여론분석 서비스인 평닷컴, 소셜분석 솔루션인 오피니언 바디가 있다. 아크원소프트는 하둡기반의 솔루션인 이지업, 알테어는 하이큐브 애널리틱스 플랫폼, 야인소프트는 옥타곤 엔터프라이즈 비즈니스 인텔리전스 서버 등을 출시하고 있다.

손진승, 최규헌[27]은 빅데이터 솔루션은 데이터 수집, 데이터 저장 및 처리, 데이터 분석, 시각화가 가능해야 한다고 했으며, 정형, 반정형, 비정형 데이터 처리 가능 여부에 솔루션을 분류하였다. 데이터 수집 기능을 갖춘 솔루션으로는 글루터 seanal.com, 오라클 Endeca, 데이터 저장 기능은 엔에프랩의 Pelto, 크루닉스의 Gridcenter Hadoop, EMC의 Greenplum DW, IBM의 Inforsphere, BigInsight, SAP의 HANA, 테라데이터의 Aster Appliance가 있다. 데이터 분석용으로는 센솔로지의 opinion buddy, pyeng.com, 이투온의 SNSpider, Altair의 HiQube, 오라클릐 Endeca, SAS의 HPA와 TextAnalytics가 있다. 시각적 표현이 가능한 솔루션으로 MS SQL Server Reporting Services, SAS Visual Analytics, Tibco Spotfire를 들었다.

### 2.2 패키지 특징

김수지, 이재희[17], 정지선[16]의 조사를 바탕으로 국내 18개, 국외 6개 업체를 대상으로 패키지가 제공하는 주요 기능과 특징을 비교하였다(<Table 1> 참조). 패키지별로 복합적인 기능을 가지고 있지만 대표적인 기준을 중심으로 분류하였다. 비교기준으로 다음 15개의 코드를 정의하였다.

코드 A는 국내, 국외구분(D: 국내, F: 국외), B는 트위터, 블로그, 페이스 북 및 소셜 미디어로부터 정보의 관계 및 패턴 분석 기능, C는 의미분석, 자연어 처리기술, 텍스트마이닝 기능, D는 하둡의 분산 병렬처리 기술의 사용유무, E는 시뮬레이션을 통한 최적화 의사결정 및 여러 가지 시나리오를 제시하는 기능이다.

코드 F는 OLAP의 다차원분석 기술 기능, G는 비즈니스 인텔리전스 및 분석, 의사결정지원 솔루션의 제공, H는 고객관계관리 기능, I는 이미지 마이닝, J는 고객의 요구사항에 맞춘 플랫폼 구축 및 컨설팅 기능이다. 코드 K는 데이터 저장장치에 중점을 두고, L은 다양한 검색이 가능한 지능형 검색엔진, 시맨틱 검색엔진을 제공하는지 여부이다. 코드 M은 데이터의 수집, 분석, 서비스의 일련의 과정을 솔루션 형태로 제공하고, N은 분석결과를 특히 시각화로 시키는 기능이 뛰어나다. 끝으로 코드 O는

비정형 데이터의 저장이 가능하고 NoSQL 데이터베이스를 사용하거나 SQL과 통합화가 가능하다.

〈Table 1〉 Comparison of big data packages

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | D | o | | | | | | o | | o | | | o | | |
| 2 | D | | | o | | | | | | o | | | | | |
| 3 | D | o | o | | | | | | | | | | o | | |
| 4 | D | o | | | | | | | | | | | | | |
| 5 | D | | | | | | | | | o | | o | | | |
| 6 | D | o | o | | | | | | | | | | | | |
| 7 | D | | | o | | | | | | | | | | | |
| 8 | D | | | | o | | o | | | | | | o | o | |
| 9 | D | | | | | o | o | | | | | | o | | |
| 10 | D | o | | | | | o | | | o | | | | | |
| 11 | D | o | o | | | | | o | | | | | | | o |
| 12 | D | | | | | | o | | | o | | | o | o | |
| 13 | D | o | | | | o | o | o | | | | | o | | |
| 14 | D | | | o | | | | | o | | | | o | | |
| 15 | D | o | | | | | | | | o | | | o | | |
| 16 | D | | o | | | | o | | | | o | | | | o |
| 17 | D | o | o | | | | | | | | | o | | | |
| 18 | D | | | o | | | | | | | o | | o | | |
| 19 | F | | | o | | | | | | o | o | | | | |
| 20 | F | | o | | | | | | | o | o | o | | | |
| 21 | F | | | | | | | | | o | | | o | | o |
| 22 | F | o | o | o | | | | | | | | | o | | o |
| 23 | F | | | | | | o | | | o | | | | o | |
| 24 | F | o | | o | | | | o | | o | | | o | | |

Note) A: domestic/foreign, B: social analysis, C: text mining, D: Hadoop and parallel processing, E: simulation, F: multi-dimensional analysis, G: business intelligence and analysis, H: customer relationship management, I: image mining, J: platform service, K: storage, L: search software, M: solution development, N: visualization, O: NoSQL

1: Gruter 2: NexR 3: Daumsoft 4: Cyram 5: Saltlux 6: Sensology 7: Archone 8: Altair 9: YainSoft 10: SM2 Networks 11: SK Telecom 12: NFLabs 13: WISE-I-Tech 14: ECMiner 15: e2On 16: Cardinal Infor. Tech 17: Konan Technology 18: Clunix 19: EMC 20: HP 21: IBM 22: Oracle 23: SAS 24: Teradata

## 3. 선정기준

빅데이터 패키지 도입시 도입 비용과 기대혜택을 통해 타당성 검토가 필요하다. 비용은 수치화할 수 있다는 점에서 측정이 용이하다. 그러나 혜택은 정성적 요소가 많아 주관적 측정이 불가피할 것이다. ERP, CRM, SCM와 같은 엔터프라이즈 정보시스템은 패키지를 사용하여 구축되므로 패키지 도입효과는 이들 정보시스템의 구현효과와 일치되는 부분이 많다. 빅데이터 패키지는 의사결정을 지원하고 조직의 성과를 가져온다는 관점에서는 엔터프라이즈 정보시스템 패키지들과 공통이지만 세부적 혜택은 차이가 있다. Hendricks, Singhal, Stratman[10]은 ERP, CRM, SCM은 재무적 성과를 가져온다고 했으며 빅데이터 분석도 재무적 성과를 기대할 수 있다. Holsapple, Sena[11]은 ERP 패키지는 비즈니스 프로세스를 개선효과를 가져온다고 했다. CRM은 충성도가 높은 고객의 유지, SCM은 물류관리의 최적화를 통한 전사적인 이익 달성 효과가 있는 반면, 빅데이터 분석은 해결하려는 문제 성격에 따라 기대 혜택은 달라질 것이다.

Jadhav, Sonar[15]는 소프트웨어 패키지의 혜택을 직접적 이익과 간접적 이익으로 구분하여 전문가가 5점 척도로 평가하는 방법을 제안하였다. 빅데이터는 타기업의 도입성공사례로부터 혜택을 추정할 수도 있다. 직접적 이익의 측정 예로, 카드회사에서 마케팅 활동으로 빅데이터 소셜분석을 하였다면 빅데이터 도입전과 후의 매출액 증감정도로 혜택이 측정가능 하다. 또한 고객들의 반응을 빅데이터 분석하여 신상품 개발에 활용한다면 신제품 효과는 정량화 할 수 있을 것이다. 간접적 이익으로 공공분야에서 재해예방이나 복지에 빅데이터를 사용하다면 그 효과는 정성적으로 측정할 수밖에 없을 것이다.

빅데이터 패키지 선정기준은 소프트웨어 패키지들의 선정기준, DSS 소프트웨어, 데이터마이닝 도구의 선정기준들, 상업용 빅데이터 패키지들이 가진 특징으로부터 도출하기로 한다. 첫째, Jadhav, Sonar[15]는 선행연구로부터 소프트웨어 패지지 선정기준으로 기능성, 품질, 벤더, 기술성, 비용, 혜택, 결과물의 출력, 이해당사자의 견해를 제안하였다. 그러나 이 결과를 적용하기에는 몇가지 문제점이 있다. 품질 기준은 개념이 광범위하기 때문에 세분화할 필요가 있고 기능성과 기술성은 중복되는 부분이 있으며, 결과물의 출력은 그다지 중요하게 여길 요소는 아닌 것 같다. 그리고 비용과 혜택은 도입타당성 조사를 위해서 별도로 관리할 필요가 있으며, 이해 당사자의 견해는 선정기준에 포함시키는 것 보다는 그룹의사

결정 모델로 다룰 필요가 있다.

본 연구에서는 빅데이터패키지 선정기준을 선행연구로부터 품질, 사용성, 벤더를 제안한다. 품질은 모든 제품 선정에서 고려되어야 할 기준이며 광범위한 개념이다. McFarland, Hamilton[18]은 품질을 시스템품질, 정보품질, 서비스품질로 구분하였다. Vankatech, Davis[30], Igbaria, Parasuraman, Baroudi[13]은 시스템품질은 시스템이 정상적으로 작동하는지 유무를 측정하는 것으로 기능, 성능, 상호작용 등으로 측정할 수 있다고 했다. Roca, Chiu, Martinez[23]는 정보품질 항목에는 직무 적합성, 이해도, 정확성, 표현형식, 최신성, 완전성, 신뢰도 등이 포함된다고 했다. Parasuraman, Zeithml, Berry[21], Zeithaml, Parasuraman, Malhotra[32]는 서비스품질은 서비스가 사용의 기대치를 충족시켜주는지 여부를 판단하는 것으로 효율성, 신뢰성, 수행도, 보안성, 고객서비스, 응답성, 보상, 접촉 정도로 측정할 수 있다

이들 연구로부터 빅데이터 패키지 특성한 적합한 품질의 세부기준으로 (1) 기능성, (2) 성능, (3)보안성을 정의한다. 기능은 빅데이터의 저장, 처리, 분석, 시각화 기능이 포함되는지 여부를 측정하고, 성능은 빅데이터가 가진 방대한 양의 데이터 처리, 다양한 데이터 유형의 처리, 처리속도가 빠른지 여부를 측정한다. 보안성이나 프라이버시 문제는 향후 빅데이터에서 중요하게 다루어야 할 부분이므로 포함시키기로 한다.

Demsar[6]은 사용성은 비정형적인 데이터마이닝 분석 도구선정에서 고려요소라고 했다. 광의의 품질개념에 포함될 수 있지만 빅데이터 처리의 전문성, 조직의 목표지향성을 고려할 때 품질과는 별개의 선정기준으로 분리하는 것이 바람직하다. Nielsen[19], ISO 9241-11[14]에 따르면 사용성은 배우기 쉽고, 사용하기 편리하고, 오류없이 신속하고 효과적으로 목표를 달성할 수 있도록 해주는 제품을 설계하는 것이다 사용성의 세부기준으로는 (1) 학습성 (2) 편의성 (3) 목표달성도를 포함시킨다.

〈Table 2〉 Characteristics of commercial packages

| | |
|---|---|
| 1 | advanced analysis for unstructured data(relation, pattern, trend, semantics, pass, similarity, statistic analysis) |
| 2 | social network analysis(SNS, blog, twitter) |
| 3 | semantics search, variety of query languages |
| 4 | integrated function for data acquisition, storage, representation |
| 5 | processing for large volume of data, real-time and batch processing |
| 6 | management, monitoring function |
| 7 | stability of system, error tolerance |
| 8 | extensibility, adding moules according to custerm requirements |
| 8 | implementation of CRM |
| 10 | visualization of reault |
| 11 | acceptance of muti-types of data |
| 12 | easy to develop models |
| 13 | applicable to various domains |
| 14 | integrated platforms |
| 15 | report management, editing function |

〈Table 3〉 Selection criteria and sub-criteria

| main | sub criteria | sub criteria |
|---|---|---|
| quality (QUA) | funtionality (FUN) | social network analysis(SNA)<br>advanced analysis(ADV)<br>stability(STA)<br>expandability(EXP)<br>monitoring(MON)<br>integrated platform(INT) |
| | performance (PER) | processing speed(PRO)<br>large volume of data storage and processing(VOL) |
| | security (SEC) | |
| usability (USA) | learnability (LEA) | |
| | ease of use (EOU) | |
| | achievement of goal (AOG) | |
| vendor (VEN) | maintenance and upgradation (MNU) | |
| | vendor reputation (VER) | |
| | consulting ability (CON) | |
| output (OUT) | ease of understanding (EAS) | |
| | visual representation (VIS) | |
| | report management (REP) | |
| | reliability (REL) | |

둘째, DSS 소프트웨어 선정기준으로 Blanc, Jelassi[2]는 기능적, 기술적 요인을 들었다. 빅데이터는 DSS 소프트웨어와 마찬가지로 모델링과 사용자 인터페이스 기능이 중요하다. (1)모델링이 얼마나 쉬운가? (2) 다양한 모델 구축이 가능한가? (3) 사용자인터페이스는 편리한가? 등을 평가할 수 있다. (1)과 (3)은 사용성 기준의 학습성 및 편의성에 포함시키고 (2)는 품질 기준의 기능에 포함시키기로 한다.

셋째, 데이터마이닝 도구는 데이터마이닝을 위한 소프트웨어로 SPSS Clementine, XL Miner, WEKA 등을 의미한다. Hung, Liu, Chang[12]은 선행연구로부터 사용의 용이성, 성능, 효과성, 처리시간, 포맷설계를 들었다. 포맷설계는 사용자와 데이터 탐색절차를 통합하는 것으로 사용자에게 피로감을 줄이고 직무수행이 쉽게 하도록 도구를 설계하는 것으로 시각화 표현을 강조하고 있다. 이는 사용성의 원칙에 부합된다. 그러나 분석결과물을 이해하기 쉽게 나타낼 수 있어야 하므로 사용성의 세부기준으로 시각화를 추가한다.

넷째, 정지선[16], 김수지, 이재희[17], 손진승, 최규헌[27]의 연구를 바탕으로 상업용 패키지들이 갖추어야 할 기능과 요건을 조사하면 <Table 2>와 같다.

선행연구로부터 <Table 3>과 같은 선정기준과 세부기준을 정의한다. 주기준에는 품질, 사용성, 벤더, 결과물이 포함된다. 품질의 세부기준에는 기능성, 성능, 보안성을, 사용성의 세부기준에는 학습성, 사용의 용이성, 목표달성도를, 벤더의 세부기준에는 유지보수와 업그레이드 능력, 벤더의 명성, 컨설팅 능력을 포함한다. 결과물의 세부기준에는 이해하기 쉬운 정도, 시각화 표현, 리포트관리, 결과물의 신뢰성을 포함한다. 기능성의 세부기준은 비교적 중요성이 높은 소셜네트워크 분석 기능, 고급분석 기능, 안정성, 확장성, 모니터닝 기능, 통합 플랫폼의 제공을 평가한다.

## 4. AHP 방법

### 4.1 손익평가

빅데이터 패키지는 그 수가 많기 때문에 AHP 모델을 적용할 경우 평가 횟수가 증가하는 문제, 사용자가 모든 패키지를 사용해보고 어떤 패키지가 나은지를 평가하는 것은 현실적으로 어려울 수 있다. 패키지는 그 특성상 TV, 컴퓨터, 자동차처럼 눈으로 살펴보거나 간단한 사용만으로 우수한 정도를 쉽게 판단하기 어렵다. 그러므로 정확한 평가를 위해서 가능한 대안의 수를 줄이는 것이 바람직하다. 타당성 평가를 위하여 비용과 혜택을 목표로 갖는 AHP 모델을 구축하여 그 비율을 측정하여 후보대안을 선정하는 방법을 제안한다. AHP 모델에서는 목표, 평가기준, 대안으로 구성되는 단일 트리를 사용하는 AHP 모델과는 달리 혜택트리와 비용트리에 속하는 평가기준은 달라질 수 있다.

비용을 평가하는 기준에는 Jadhav, Sonar[15]가 제안한 라이센스 비용, 하드웨어 구입비, 소프트웨어 구입비, 설치비, 유지보수비, 교육훈련비, 업그레딩 비용을 포함시킨다. 혜택은 직접적 이익과 간접적 이익을 포함시킨다. 비용의 기준들과 직접적 이익은 수치화 할 수 있고, 간접적 이익은 5점 척도로 평가한다.

유의한 대안을 선정하기 위하여 혜택트리와 비용트리를 구축한 다음, 쌍비교 방식에 의하여 각 평가기준의 중요도와 대안의 우선순위를 구한다. 손익비율의 값이 1보다 큰 대안을 선정한다.

손익비율의 계산 절차는 어떤 패키지 i를 선택했을 때 기대이익을 Bi, 비용을 Ci라고 하면 손익비율 Ri는 다음과 같다.

$$R_i = \frac{B_i}{C_i}$$

소비자가 평가한 혜택과 비용의 주기준의 중요도를 각각 aj, bj라 할때  대안의 중요도에 주기준의 중요도를 가중평균하면, 수정된 손익비율 WRi는 다음과 같다 나타낼 수 있다.

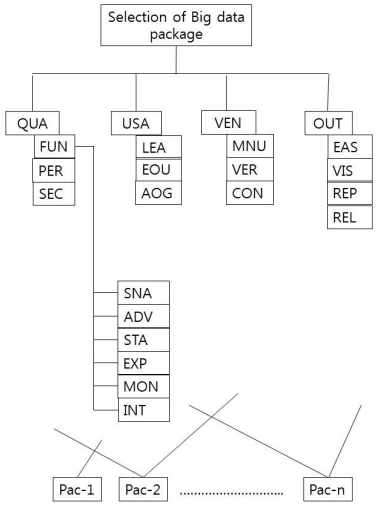
$$WR_i = \frac{\sum_{j=1}^{k} a_{.j} B_{ij}}{\sum_{j=1}^{l} b_{.j} C_{ij}}$$

그 다음 대안 전체에 대한 WRi는 각각의 WRi의 합으로 나타낼 수 있고, 만약 이 값이 1보다 크다면 혜택이 비용보다 크므로 대안은 채택된다.

### 4.2 AHP 모델

<Table 3>의 선정기준을 계층적 구조로 나타내면 [Fig 1]과 같다. [Fig 1]의 계층구조는 목표노드와 선정기

준, 각 선정기준별 세부기준, 그리고 평가 대안으로 구성된다. 선정기준의 중요도를 결정하기 위한 의사결정자는 빅데이터 분석가나 분석가, 사용자, 프로젝트 매니저, CEO 등이 참여하는 평가 팀을 구성하는 방식이 바람직하다. 선정기준의 가중치는 쌍비교로 도출할 수 있다. 그러나 선정기준에 대한 패키지들의 중요도는 패키지들간의 쌍비교 보다는 각 페키지를 5점 척도로 평가하는 것이 바람직하다. 왜냐하면 사용자는 패키지의 수가 많다는 점과 실제 모든 패키지를 사용할 기회가 제공되지 않는다는 점, 그리고 한 패키지를 사용한 후 다른 패키지를 사용하는 동안 전에 사용했던 패키지에 대한 기억이 소멸될 수 있기 때문이다. 그러므로 특정 패키지를 경험한 후에 선정기준에 대한 자신의 느낌을 5점 척도(매우 좋음, 좋음, 보통, 나쁨, 매우 나쁨)로 기록한다. 이러한 방식은 각 패키지의 사용자가 다를 경우에도 가능하다. 한 사용자가 모든 패키지를 경험할 수 없을 경우 전문가들의 평가치를 사용할 수도 있다. 각 패키지를 평가한 스코어보드를 활용할 수 있다.




[Fig 1] AHP model

### 4.3 적용 예제

평가 대상인 패키지를 4개라고 가정할 때 선정과정을 수치적 예를 통해 보이기로 한다. 손익분석은 다음과 같은 절차를 따른다.

(1) 비용기준을 쌍비교를 통해 가중치를 도출한다.
(2) 각 패키지별 선정기준에 대한 비용을 산출한다.
(3) 비용기준의 중요도에 비용 값을 곱한다.
(4) 비용기준에 대한 로컬 우선순위를 계산한다.
(5) 로컬 우선순위의 합은 1이 되도록 정규화 한다.
(6) 직접 혜택, 간접혜택의 가중치를 도출한다.
(7) 직접 혜택의 값, 간접 혜택의 값을 구한다. 직접 혜택은 수치화하고, 간접혜택은 5점 척도로 평가한다.
(8) 혜택의 가중치와 혜택의 값을 곱하여 로컬우선순위를 계산한다.
(9) 로컬우선순위의 합이 1이 되도록 정규화 한다.
(10) 혜택/비용의 비율을 계산한다.
(11) 혜택/비용의 값이 1보다 큰 패키지를 선정한다.

<Table 4>는 일련의 절차를 계산한 예이다. 각 패키지별 비용의 우선순위는 각각 0.29, 0.21, 0.26, 0.24로 계산된다. 비용은 패키지-2가 가장 작고, 패키지-1이 가장 크다. 직접적 혜택과 간접적 혜택의 우선순위는 각각 0.32, 0.14, 0.13, 0.41로 계산된다. 패키지 4가 가장 좋고, 패키지 3이 가장 나쁘다. 혜택/비용의 비율을 계산하면 각각 1.12, 0.67, 0.48, 1.68로 계산되고 이 값이 1보다 큰 패키지는 패키지-1, 패키지-4가 된다. 즉 손익분석에서 2개의 패티지가 선택되며 패키지-1, 4에 대해서 선정기준을 평가하여 더 나은 패키지를 선택하는 절차를 수행한다.

〈Table 4〉 cost/benefit analysis

| | | weight | PAC-1 | PAC-2 | PAC-3 | PAC-4 |
|---|---|---|---|---|---|---|
| Co-st | LI | 0.17 | 0.04 | 0.04 | 0.09 | 0.13 |
| | H/W | 0.16 | 0.16 | 0.09 | 0.18 | 0.15 |
| | S/W | 0.24 | 0.34 | 0.23 | 0.21 | 0.16 |
| | IN | 0.13 | 0.23 | 0.13 | 0.14 | 0.25 |
| | MA | 0.08 | 0.42 | 0.21 | 0.22 | 0.19 |
| | ET | 0.13 | 0.06 | 0.13 | 0.12 | 0.09 |
| | UP | 0.09 | 0.09 | 0.17 | 0.32 | 0.24 |
| | overall | 0.29 | 0.21 | 0.26 | 0.24 | |
| Direct benefit | | 0.6 | 0.13 | 0.02 | 0.02 | 0.23 |
| Indirect benefit | | 0.4 | 0.19 | 0.14 | 0.12 | 0.15 |
| | overall | 0.32 | 0.14 | 0.13 | 0.41 | |
| | B/C ratio | 1.12 | 0.67 | 0.48 | 1.68 | |

Note) LI: license, IN: installation, MA: maintenance, ET: education and training, UP: upgrading, PAC: package

〈Table 5〉 priority of packages

| | | | | Pac-1 | Pac-4 | Pac-1 | Pac-4 |
|---|---|---|---|---|---|---|---|
| QUA | 0.2 | | | | | | |
| | FUN | 0.51 | | rating | | priority | |
| | | SNA | 0.32 | 5 | 4 | 0.56 | 0.44 |
| | | ADV | 0.21 | 4 | 5 | 0.44 | 0.56 |
| | | STA | 0.23 | 1 | 3 | 0.25 | 0.75 |
| | | EXP | 0.08 | 2 | 2 | 0.50 | 0.50 |
| | | MON | 0.09 | 3 | 2 | 0.60 | 0.40 |
| | | INT | 0.07 | 3 | 4 | 0.43 | 0.57 |
| | PER | | 0.34 | 3 | 3 | 0.50 | 0.50 |
| | SEC | | 0.15 | 4 | 5 | 0.44 | 0.56 |
| USA | 0.3 | | | | | | |
| | LEA | | 0.31 | 3 | 1 | 0.75 | 0.25 |
| | EOU | | 0.43 | 4 | 2 | 0.67 | 0.33 |
| | AOG | | 0.26 | 4 | 5 | 0.44 | 0.56 |
| VEN | 0.2 | | | | | | |
| | MNU | | 0.41 | 2 | 4 | 0.33 | 0.67 |
| | VER | | 0.25 | 5 | 1 | 0.83 | 0.17 |
| | CON | | 0.34 | 3 | 3 | 0.50 | 0.50 |
| OUT | 0.3 | | | | | | |
| | EAS | | 0.18 | 3 | 1 | 0.75 | 0.25 |
| | VIS | | 0.23 | 4 | 5 | 0.44 | 0.56 |
| | REP | | 0.33 | 5 | 2 | 0.71 | 0.29 |
| | REL | | 0.26 | 3 | 2 | 0.60 | 0.40 |
| | | | | | overall | 0.54 | 0.46 |

[Fig 1]의 AHP 모델은 다음과 같은 절차를 따라 구현한다.

(1) 선정기준, 세부기준을 쌍비교하여 가중치를 도출한다.
(2) 세부기준에 대하여 패키지의 만족도를 5점 척도로 평가한다.
(3) 세부기준의 가중치와 패키지의 중요도를 곱한다.
(4) 세부기준에 대하여 패키지의 우선순위를 정규화한다.
(5) 각 세부기준에 대하여 패키지의 우선순위를 합계하여 전체 패키지의 우선순위를 계산한다.

<Table 5>에서 4개 주기준의 가중치는 각각 0.2, 0.3, 0.2, 0.4이다. 패키지-1, 4에 대하여 세부기준을 5점 척도로 평가한다. 예를 들어, SNA(소셜네트워크분석) 기준을 가정할 때 패키지-1은 제공되지 않고, 패키지-4는 제공된다고 가정할 때, 패키지-1의 값은 1을 부여하고, 패키지-4는 3, 4, 5 중에서 기능이 매우 우수하면 5, 우수하면 4, 보통정도면 3을 부여한다. 그 다음 전체 우선순위는 패키지-1은 0.54, 패키지-2는 0.46으로 계산되어(열의 합계) 패키지-1이 최종 선택된다.

## 5. 결론

빅데이터 분석은 빅데이터 패키지로만 가능하다. 빅데이터 패키지 선정문제는 많은 패키지 중에서 사용하기 쉽고, 품질이 좋고, 모델링 능력이 뛰어난 저가의 패키지를 판별하는 문제이다. 본 연구에서는 상업용 빅데이터 패키지 현황과 벤더들을 고찰하고 패키지 선정기준을 제안하였다. AHP 모델은 의사결정자의 주관적인 판단치를 정량화하며 각 패키지별 점수를 도출할 수 있다는 점에서 의사결정 과정이 이해하기 쉽고 패키지 간 순위를 도출할 수 있다는 것이 장점이다. 그리고 패키지 도입시 손익에 대한 비율분석은 도입 타당성을 평가하고 평가대안 수를 줄일 수 있다. 많은 IT기업들이 다양한 기능을 가진 빅데이터 패키지를 출시하고 있기 때문에 본 연구에서 제안한 방법은 의사결정자 스스로 부여한 선정기준의 가중치와 전문가들이 각 패키지별에 대한 평가점수를 결합하여 빅데이터 패키지들을 평가할 수 있을 것이다.

향후 과제로는 각 패키지에 대한 전문가들의 벤치마킹 데이터를 획득하여 실증분석을 수행하는 일이다. 그리고 혜택을 정량화라는 일과 분석가가 아닌 태스크포스팀의 중요도를 반영하여 그룹의사결정 문제로 확장하는 일이다.

## ACKNOWLEDGMENTS

This research publication was supported by Kyungsung University Research Grants in 2013

## REFERENCES

[1] Alshawi, S., Missi, F., and Irani, Z., Organizational, technical and data quality factors in CRM

adoption-SEM perspective. Industrial Marketing Management, Vol. 40, No. 3, pp. 376~383, 2011.

[2] Blanc, L. A. and Jelassi, M. T., DSS software selection: A multiple criteria decision methodology. Information & Management, Vol. 17, No. 1, pp. 49~65, 1989.

[3] Brown, B., Chui, M., and Manyika, J., Are you ready for the era of big data? Parsing the benefits: not all industries are created equal. The McKinsey Quarterly, Vol. 4, 24~35, 2011.

[4] Brynjolffson, E., Hammerbacher, J., and Stevens, B., Competing through data: three experts offer their game plans. The McKinsey Quarterly, Vol. 4, pp. 36~47, 2011.

[5] Chen, H., Chiang, R. H. L., and Storey, V. C., Business intelligence and analytics: from big data to big impact. MIS Quarterly, Vol. 36, No. 4, pp. 1165~1188, 2012.

[6] Demsar, U., Investigating visual exploration of geospatial data: An exploratory usability experiment for visual data mining. Computer, Environment and Urban Systems, Vol. 31, 551~571, 2007.

[7] E-Today, Five factor for leading US economy. 19 July, 2013.

[8] ETnews, Korean Big data market trend. Vol. 23 April, 2013.

[9] Gürbüz, T. S., Alptekin, S. E., Alptekin, G. I., A hybrid MCDM methodology for ERP selection problem with interacting criteria. Decision Support Systems, Vol. 54, No. 1, pp. 206~214, 2012.

[10] Hendricks, K. B., Singhal, V. R., and Stratman, J. K., The impact of enterprise systems on corporate performance: A study of ERP, SCM, and CRM system implementations. Journal of Operations Management, Vol. 25, No. 1, pp. 65~82, 2007.

[11] Holsapple, C. W. and Sena, M. P., ERP plans and decision-support benefits. Decision Support Systems, Vol. 38, No. 4, pp. 575~590, January 2005.

[12] Huang, T. C. K., Liu, C. C., and Chang, D. C., An empirical investigation of factors influencing the adoption of data mining tools. International Journal of Information Management, Vol. 32, pp. 257~270, 2012.

[13] Igbaria, M., Parasuraman, S., and Baroudi, J. J., A motivational model of microcomputer usage. Journal of Management Information Systems, Vol. 13, No. 1, pp. 127~143, 1996.

[14] ISO 9241-11, Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability. International Organization for Standardization, 1998.

[15] Jadhav, A. S. and Sonar, R. M., Framework for evaluation and selection of the software packages: A hybrid knowledge based system approach. The Journal of Systems and Software, Vol. 84, pp. 1394~1407, 2011.

[16] Jeong, J. S., Status of Big data solutions and services-I. National Information Society Agency, September, pp. 1~84, 2012.

[17] Kim, S. J. and Lee. J. H., Status of Big data solutions and services-II. National Information Society Agency, December, pp. 1~110, 2012.

[18] McFarland, D. J. and Hamilton, D., Adding contextual specificity to the technology acceptance model. Computers in Human Behavior, Vol. 22, No. 2, pp. 427-447, 2006.

[19] Nielsen, J., Usability Engineering. Academic Press, SanDiego, 1993.

[20] Oracle, Oracle NoSQL database. An Oracle white paper, September 2011.

[21] Parasuraman, A., Zeithml, A., and Berry, L. L., A conceptual model of service quality and its implications for future research. Journal of Marketing, Vol. 49, pp. 41~50, 1985.

[22] Pomerol, J. C., Romero, S. B., Multicriterion Decision in Management: Principles and Practice. Kluwer Academic Publishers, Norwell, 2000.

[23] Roca, J. C., Chiu, C. M., and Martinez, F. J., Understanding e-learning continuance intention: An extension of the Technology Acceptance Model. Int. J. Human-Computer Studies, Vol. 64, pp. 683~696, 2006.

[24] Russom, P., Big data analytics. Fourth Quater, TDWI Research, 2011.
[25] Saaty, T. L. and Kearns, K., Analytical Planning: The Organization of Systems. Pergamon Press, Oxford, 1985.
[26] Seng, J. L. and Chen, T. C., An analytic approach to select data mining for business decision. Expert Systems with Applications, Vol. 37, No. 12, pp. 8042~8057, 2010.
[27] Son, J. S. and Choi, K. H., Trend of Big data solution and implication. Weekly Technology Report, National IT Industry Promotion Agency, April 17, 2013.
[28] Stonebraker, M., Abadi, D., DeWitt, D. J., Madden, S., Pavlo, A., and Rasin, A., MapReduce and parallel DBMSs: Friendsor foes. Communications of the ACM, Vol. 53, No. 1, pp. 64~71, 2012.
[29] Venables, W. N., Smith, D. M., and R Core Team, An Introduction to R. R Core Team, 2013.
[30] Venkatesh, V. and Davis, F. D., A model of the antecedents of perceived ease of use: Development and test. Decision Sciences, Vol. 27, No. 3, pp. 451~481, 1996.
[31] White, T., Hadoop: The Definitive Guide. O'Reilly Yahoo Press, 2009.
[32] Zeithaml, V. A., Parasuraman, A., and Malhotra, A., Service quality delivery through Web sites: A critical review of extant knowledge. Journal of the Academy of Marketing Science, Vol. 30, pp. 362~375, 2002.

**변 대 호(Byun, Dae Ho)**

・1985년 2월 : 고려대학교 산업공학과 (공학사)
・1987년 2월 : KAIST 산업공학과 (공학석사)
・1996년 2월 : POSTECH 산업공학과 (공학박사)
・1996년 3월 ~ 현재 : 경성대학교 경제금융물류학부 교수
・관심분야 : IT미디어 UX 평가
・E-Mail : dhbyun@ks.ac.kr