https://doi.org/10.14400/JDC.2017.15.6.133

정형 비정형 빅데이터의 융합분석을 위한 소비 트랜드 플랫폼 개발

김성현*, 장석호**, 이상원***
한국정보화진흥원 K-ICT빅데이터센터*, 비씨카드 빅데이터센터**, 원광대학교 컴퓨터소프트웨어공학과***

Consumer Trend Platform Development for Combination Analysis of Structured and Unstructured Big Data

Sunghyun Kim*, Sokho Chang, Sangwon Lee*****
K-ICT Big Data Center, Korea Information Society Agency*
Big Data Center, BC Card**
Department of Computer Software Engineering, Wonkwang University***

요 약 데이터는 금융업에서 가장 중요한 자산으로 평균 71%의 금융기관이 데이터 분석으로 경쟁우위를 창출하고 있다. 특히, 금융업 중 카드 업종에서는 전체 고객의 소비행위 패턴 및 선호 트랜드 분석에 의한 가맹점 정보, 경기 변동 상황, 상권정보 제공 서비스 개발에 빅데이터가 폭 넓게 활용되고 있지만 데이터의 융복합을 통한 새로운 가치 창출은 미흡한 편이다. 본 연구는 소셜 데이터와 BC 카드 매출데이터의 융합 분석한 신용카드 회사의 '소비 트랜드 분석 및 예측' 사례를 다룬다. BC카드는 소셜 데이터를 활용한 트랜드 프로파일링 작업과 카드 및 소셜 데이터를 연계하는 알고리즘 개발 및 분석 내용 시각화 시스템을 개발하였다. 성과 검증을 위해 '식스포켓' 관련 트랜드를 분석하고 마케팅을 시행해 본 결과 40~100%이상의 마케팅 승수 증대 효과를 거두었다. 본 연구는 그동안 개별적으로 이루어져 오던 정형, 비정형데이터 분석을 융합하여 분석하는 방법론과 사례를 창출한 의의가 있으며 이는 앞으로 카드 업종 뿐만 아니라 타 업종에도 변화하는 트랜드에 유용하게 대응할 수 있는 시사점을 제공할 것이다.

주제어 : 빅데이터, 연관분석, 소셜 네트워크 서비스, 마케팅 오퍼, 비구조적 데이터

Abstract Data is the most important asset in the financial sector. On average, 71 percent of financial institutions generate competitive advantage over data analysis. In particular, in the card industry, the card transaction data is widely used in the development of merchant information, economic fluctuations, and information services by analyzing patterns of consumer behavior and preference trends of all customers. However, creation of new value through fusion of data is insufficient. This study introduces the analysis and forecasting of consumption trends of credit card companies which convergently analyzed the social data and the sales data of the company's own. BC Card developed an algorithm for linking card and social data with trend profiling, and developed a visualization system for analysis contents. In order to verify the performance, BC card analyzed the trends related to 'Six Pocket' and conducted th pilot marketing campaign. As a result, they increased marketing multiplier by 40~100%. This study has implications for creating a methodology and case for analyzing the convergence of structured and unstructured data analysis that have been done separately in the past. This will provide useful implications for future trends not only in card industry but also in other industries.

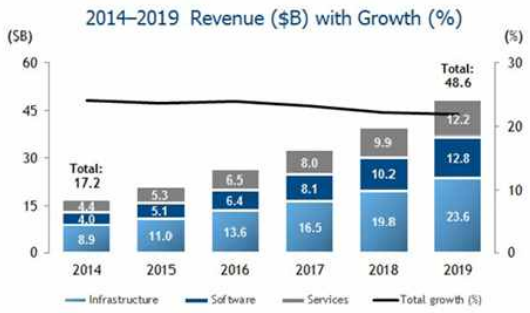
Key Words : Big Data, Association Analysis, Social Network Service, Marketing Offer, Unstructured Data

* This paper was supported by Wonkwang University in 2017.
Received 8 April 2017, Revised 4 May 2017
Accepted 20 June 2017, Published 28 June 2017
Corresponding Author: Sangwon Lee (Department of Computer Software Engineering, Institute of Convergence and Creativity, Artificial Intelligence Research Institute of Society, Wonkwang University)
Email: sangwonlee@wku.ac.kr
ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

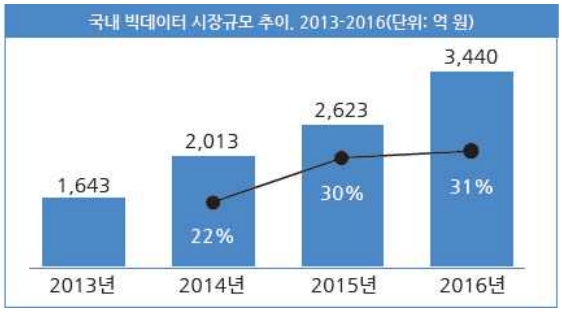
1. Introduction

빅데이터는 일반적인 데이터의 관리 한계를 넘어서는 방대한 양(Volume)을 가진 데이터로 빠른 생성 속도(Velocity), 종류의 다양성(Variety)으로 정의 된다 [1]. 이제 빅데이터는 데이터 관리 도구, 플랫폼, 분석기법까지 포괄하며 가치창출 및 활용에 초점을 두고 개념이 진화하고 있다 [2]. 네트워크 사용과 데이터 저장 비용의 급격한 하락, 스마트폰과 소셜네트워크서비스의 확산, 사물인터넷(Internet of Things)과 M2M(Machine to Machine) 보급이 빅데이터 시대를 도래한 배경이다 [3]. 데이터 수집, 저장하는 하드웨어와 분석에 필요한 소프트웨어 기술이 발달하면서 대용량의 데이터를 빠르게 분석하여 기존에는 불가능했던 새로운 형태의 서비스와 인싸이트 도출이 가능해졌으며, 이로 인해 빅데이터의 중요성은 날로 증대되고 있는 추세다 [4]. '데이터'는 정보화 시대의 원유로 비유되는 미래 핵심 자원으로 산업화 시대의 철과 석탄, 정보화 시대의 정보시스템처럼 다가올 지능정보사회에서는 빅데이터가 사회발전의 핵심자원이 될 것이다 [4]. 기존 정보시스템과 센서, 스마트 폰 등에서 축적된 데이터를 분석할 수 있는 역량은 기업의 경쟁력과 직결된다. 인공지능 기법을 도입한 자동번역시스템, SNS 분석을 통한 고객인식 파악 등 과거에는 불가능했던 많은 일들이 빅데이터를 통해 현실화되고 있다.



[Fig. 1] The World Big Data Market Size

IDC에 의하면 전세계 시장 연평균 성장률(2014-2019)은 23.1%로 지속적인 성장세에 있으며 2019년 486억달러에 달할 것으로 예측 된다 [5] [Fig. 1]. 이기간 동안의 연평균 성장률은 데이터 인프라(21.7%), 빅데이터 소프트웨어(26.2%), 빅데이터 서비스(22.7%)로 예측된다. 한국 빅데이터 시장은 2016년 약 3,440억원 규모로 전년대비 31% 성장하였으며 향후에도 30% 이상의 매우 빠른 속도로 성장할 것으로 예측 된다 [6][Fig. 2].



[Fig. 2] The Korean Big Data Market Size

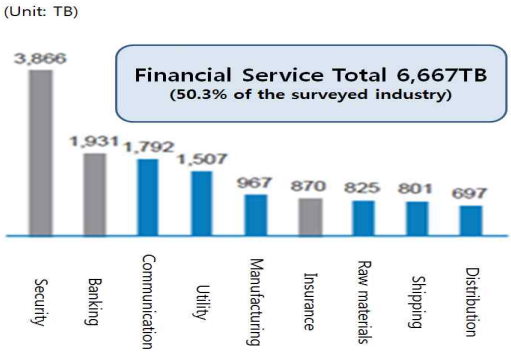
정보기술은 제조업 등 국가 기간산업에서 생산성을 높이고 있다 [7] 경영기법 및 새로운 경영도구의 도입은 지속적으로 기업의 성과를 높이고 있다 [8, 9]. 또한 클라우드 컴퓨팅과 같은 최신 정보기술의 도입은 기업의 경쟁우위를 높이고 있다 [10, 11]. 정보기술 및 시스템의 도입이 경영 성과에 긍정적인 영향을 주는 반면, 정보시스템의 근간이 되는 데이터에 대한 관심은 높지 않은 편이다. 데이터의 중요성이 강조되고 있는 반면, 데이터를 활용하여 가치를 창출해내는 구체적인 사례는 아직 충분히 제시되고 있지 못하다. 특히 데이터의 축적과 활용이 상대적으로 용이한 의료, 금융 등의 산업에서 빅데이터 시대에 걸맞는 다양한 데이터를 융복합하여 분석한 사례는 그 시도에 비해 성공사례가 아직 널리 알려져 있지 못하다 [12]. 본 연구는 금융분야에서 정형, 비정형 데이터의 융합 분석을 통해 고객가치를 창출하여 매출을 향상한 실질적인 사례를 제시하여 빅데이터의 활발한 활용하는데 목적이 있다.

2. 이론적 배경

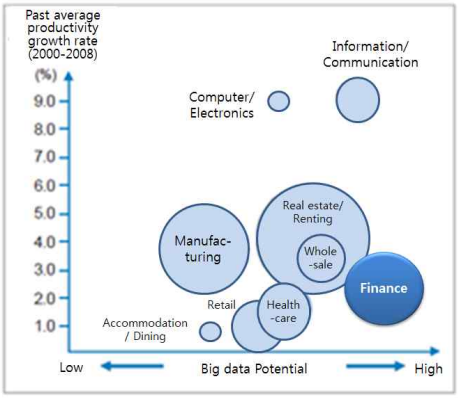
2.1 산업 동향

산업계에서는 빅데이터가 21세기 지식정보화 사회에 있어서 혁신을 주도할 주요한 핵심 자원이자 기술이 될 것으로 예상하고 있다. 18세기 증기기관, 19세기 전기, 20세기 석유·천연가스가 산업 발전에 큰 역할을 했다면 21

세기에는 데이터가 경제와 사회를 혁신할 동력으로 보고 있다. 정보기술은 구글, 애플, 아마존 등 글로벌 선도기업들은 빅데이터를 활용한 분석·예측 역량 확보에 주력하고 있다 [4]. 금융업은 대부분의 업무처리가 온라인 환경에서 이루어져 내부 데이터의 유입 및 축적 정도가 매우 높다는 특징을 가지고 있다. 또한 처리결과의 정확성, 신뢰성이 고도로 요구되어 타 산업 군에 비해 잠재가치와 효과적 활용 가능성이 매우 높다 [Fig. 3, 4] [13]. 제조업과 같이 물질적 기반이 없는 상태에서 데이터는 금융업에서 가장 중요한 자산이 되며 평균 71%의 금융기관이 데이터 분석으로 경쟁우위를 창출하고 있다. 이는 일반기업의 63%에 비해 높은 수준이다 [14].



[Fig. 3] The Average Amount of Corp. Data(U.S.)



[Fig. 4] The Potential Value of Big Data (The size of circle shows the industry contribution of GDP)

한국정보화진흥원에 의하면 금융업의 빅데이터 투자 규모는 2016년에는 35.3%, 3년 후인 2018년에도 39.2%의 높은 성장률을 보일 것으로 기대된다. 금융은 빅데이터 시장을 선도하는 업종으로 금융관련 리스크 감소, 이상감지 시스템 등의 빠른 데이터 처리 수요 및 고객관리와 관련 마케팅 수요 증가, 투자 분석 알고리즘 고도화 등 지속적인 수요 확대가 기대된다 [15]. 최근 핀테크의 활발한 적용으로 인해 금융업에서는 ICT의 적용이 활발한 편이며, 금융업 중 카드 업종에서는 카드데이트를 활용한 경기 예측 등 거시적인 분야에서도 빅데이터는 폭넓게 시도되고 있다 [16, 17]. 개인 고객별 소비행위 패턴 및 선호 트랜드 분석에 의해 개인별 맞춤 카드 개발 및 서비스 개발 등에 빅데이터가 활용되고 있으며, 전체 고객의 소비행위 패턴 및 선호 트랜드 분석에 의한 가맹점 정보, 경기 변동 상황, 상권정보 제공 서비스 개발에도 빅데이터가 폭 넓게 활용되고 있다 [13].

신한카드에서는 2,200만 고객의 카드 사용실적을 토대로 고객별 소비패턴과 선호 트랜드를 분석하여 남녀 각각 9개 고객군을 추출한 후, 각각의 고객 유형에 최적화된 코드나인(Code9) 카드 시리즈를 출시하였다. 최근에 신한카드는 고객의 소비패턴에 맞는 비용관리항목을 인공지능(AI) 알고리즘으로 자동 분류해주고, 고객이 비용관리 항목별로 예산 설정 시 예산 내에서 소비할 수 있도록 조언 해주는 소비 비서 서비스 판(FAN) 페이봇를 출시하는 등 빅데이터를 고객서비스에 적극 활용하고 있다. 이 밖에도 공공부문과 협업하여 신한카드는 지자체의 관광활성화 정책, 교통 인프라 개선, 상권 활성화 정책, 복지 정책 최적화, 체감 경기지수 개발에 빅데이터를 활용하고 있다 [18]. 현대카드는 매출과 고객데이터를 분석하여 가맹점에 제공하는 빅데이터 활용 마케팅 지원에 집중하고 있다. 카드결제 정보를 활용해 소비자 트랜드 및 경기 변동 상황을 분석해주는 '현대카드 X 빅데이터' 리포트와 '마이 비즈니스(MY BUSINESS)' 서비스를 통해 단순 업종별 매출 비교를 넘어 주제 분야를 집중 분석해 자영업자에게 고객의 성별, 연령대, 직업, 재방문율, 보유카드혜택 등 실제 유용한 의사결정 지원 정보를 제공하고 있다 [19].

2.2 정형 데이터와 비정형 데이터

빅데이터의 개념은 가트너의 Douglas Laney가 3V(Volume, Variety, and Velocity) 개념을 제시한 후 IBM에서 진실성(Veracity)을, SAS에서는 Variability(복잡성), 오라클에서 Value(가치)의 개념을 덧붙여 가며 발

전해 왔다 [12]. 빅데이터는 기존 정형 데이터의 분석에서 비정형데이터로 분석대상을 옮겨가며 발전하고 있다. 정형데이터는 지정된 필드에 저장되어 컴퓨터 장비에 의해 바로 접근이 가능한 데이터를 의미하고, 비정형 혹은 반정형데이터는 고객의 이메일이나 웹페이지 같이 문서와 같은 형태로 저장된 자료를 의미한다 [21].

2.3 선행연구

빅데이터와 이의 응용을 포괄적으로 다룬 리뷰 논문으로는 Chen 과 Syed, Gandomi의 연구가 있다 [20, 22, 23]. Chen은 분석데이터의 특성에 따라 Business intelligence and analytics (BI&A) 1.0에서부터 3.0까지로 나누고 이의 응용분야를 제시하였다 [20]. Syed는 빅데이터로 인한 마케팅, 환경, 보건, 소셜미디어, 과학 분야에서의 혁신과 차별화, 성장 기회, 기술적 트랜드를 정리 하였다 [22]. Gandomi는 빅데이터의 개념과 데이터의 성격, 방법론과 분석기법에 대한 포괄적인 내용을 제시하였다 [23]. 빅데이터 적용사례, 그 중에서도 정형 및 비정형 데이터의 융합 분석을 다룬 연구는 매우 드문 편인데 Baars는 외부의 보고서 형태의 데이터를 내부 데이터와 결합하여 비즈니스 인텔리전스 측면에서 활용할 수 있는 방안을 제시하였다 [21]. Shin은 빅데이터의 역량을 체계적으로 측정하기 위한 연구를 실시하였다. [24]. Lee는 빅데이터 적용에 적합한 산업을 AHP 방식을 통해 탐색하였으며 [25], Kim은 공공부문에 빅데이터와 같은 신기술 도입 타당성을 평가하기 위한 프레임워크를 제안하였으며, Jung은 빅데이터의 도입 효과를 분석하여 성공요인을 도출하였다 [26 27]. 빅데이터는 공공 정책을 수립하고 이행하기 위한 유용한 수단이 되기도 하는데, Noh는 대학생들의 인식을 빅데이터 분석하여 교육 정책을 제시한 바 있으며, Kim은 서울시의 빅데이터를 통한 재난관리 정책을 제시하였다 [28, 29]. Choi는 지방정부의 범죄예방정책을 빅데이터 분석을 통해 제시한 바 있으며, Noh는 전반적인 지방정책에 빅데이터를 도입하는 방안을 제시하였다 [30. 31].

빅데이터의 산업적 활용을 위해 Noh는 Manufacturing Execution System에서의 빅데이터 도입방안을, Kim은 중소 제조기업의 4M(Man, Machine, Media, Management) 데이터 분석을 통한 제조공정 효율화 사례를 제시하였다 [32, 33]. 소셜 및 텍스트 데이터 분석사례로는 Gu의 특허 텍스트 데이터분석을 통한 전기차량기술 트랜드 분석이 있으며, Park은 소셜네트워크 분석을 위한 연구방법론을 제시한 바 있다 [34, 35]

이 밖에도 빅데이터 활성화를 위한 정책 연구나 빅데이터 기획 접근법에 대한 연구도 있으나, 빅데이터의 본격적인 비즈니스 적용과 성과 제시, 빅데이터 시대를 본격적으로 도래하게 한 비정형데이터와 정형데이터의 융합분석은 아직 미흡한 편이다 [36, 37].

3. 소비트랜드 융합분석 연구

3.1 분석 목적 및 개요

본 연구는 갈수록 경쟁이 격화되고 있는 국내 신용카드 업종에서 새로운 서비스 제공과 내부 경쟁력 향상을 위한 과제로 소셜 데이터와 카드 매출데이터의 융합 분석을 이루어낸 BC카드의 '소비트랜드 분석 및 예측' 사례를 다룬다. 본 사례연구의 '소비트랜드 분석 및 예측'은 소셜 데이터를 활용한 트랜드 프로파일링 작업과 카드 및 소셜 데이터를 연계하는 알고리즘 개발 및 분석 내용 시각화를 통해 마케팅 인사이트를 도출하는 새로운 마케팅 기법이다. BC카드는 정형 및 비정형 빅데이터 분석을 통해 소비트랜드를 분석하고 예측하여 카드사에 실제적인 도움이 되는 마케팅을 실행하기 위해 연구를 추진하였다. 연구 기간은 2015년 6월부터 12월이였으며, 데이터는 정형 데이터로는 BC카드의 카드매출 데이터, 비정형데이터로는 LG CNS가 보유한 소셜 데이터를 활용하였다.

3.2 분석 목표 및 배경

본 사례는 국내에서 최초로 비정형 데이터, 카드 데이터와 소셜 데이터를 연계 분석하는 사례로 소셜 빅데이터를 활용하여 카드사 마케팅에 성공 사례를 창출, 동종업계에 이를 전파하고 확산하여 빅데이터 저변을 넓히는 데도 목적을 두었다. 기존에도 카드 데이터에 대한 데이터 분석은 많이 진행되어 왔다. 카드 및 고객 데이터를 마이닝하여 새로운 카드상품을 개발하거나 마케팅 캠페인을 수행하는 경우는 가장 일반적인 데이터 분석의 사례이고, 소셜데이터를 분석하는 경우도 카드사의 평판을 조사하고 이를 홍보 및 마케팅에 활용하는 경우에 그쳤

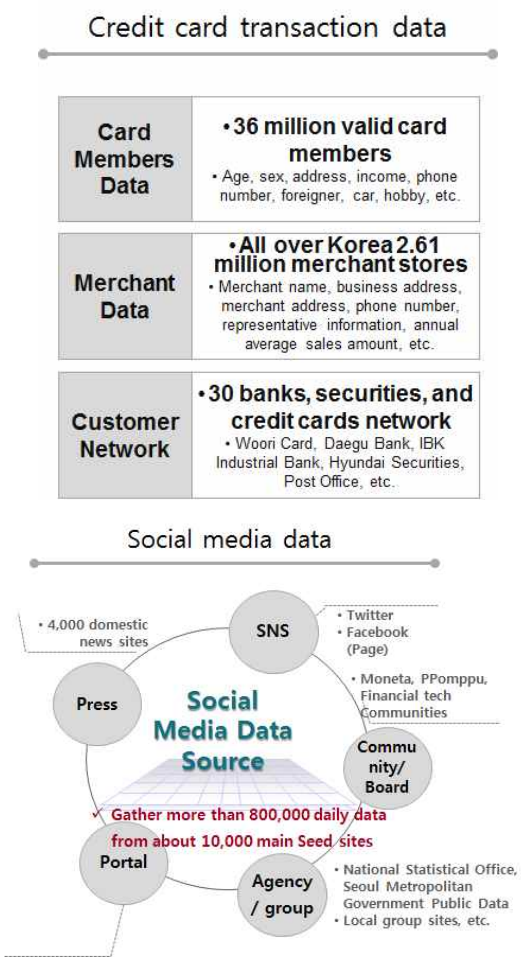
다.

기존의 마케팅은 담당자의 감과 경험에 의존하는 경향이 있어 새로운 마케팅 기획 할 때 마케팅 키워드 선정 어려움이 있었고, 분석 툴이나 프로세스 없이 진행하는 트랜드 조사에 마케팅 추진방향이 결정되어 마케팅 결과에 편차가 컸다. 또한, 다양한 소비성향의 증가로 사회적 변동성에 대한 후행지표인 카드 매출데이터 분석만으로 소비자를 예측하기 어려운 점이 있었다. 기존의 카드 매출데이터 분석은 내부 고객의 데이터만 분석한다는 한계점을 가지고 있었고, 한 가지 특정한 성향에 머물지 않고 시기별, 장소별로 다양한 소비유형을 보이는 소비성향을 모두 포괄하기 힘들었다. 특정 시점 기준으로 카드데이터를 분석하고 실제 마케팅 하는 것은 최소 6개월에서 최대 1년 전의 소비 트랜드를 반영하게 된다는 점을 고려하면 적시성 있는 마케팅 캠페인의 기획이 어려운 것임을 알 수 있다.

본 연구에서 BC 카드는 정형 데이터인 카드데이터와 비정형 데이터인 소셜 데이터를 연결 할 수 있는 키(Key)값을 고안하고, 이를 바탕으로 모양이 다른 두가지 데이터를 연계할 수 있는 알고리즘기반 마케팅 시스템(AIPS, Artificial Intelligence Marketing Profiling System)을 개발해 실제 시중은행의 카드 마케팅에 적용하여 기존 마케팅의 한계를 극복하고 새로운 형태의 마케팅 기획 방법론을 도출하는데 의의를 두었다. 기존에도 카드사별로 매출데이터를 분석하거나 소셜 트랜드를 기반으로 카드 출시를 하는 경우는 있었으나, 정형데이터인 매출데이터와 비정형데이터인 소셜데이터를 융합하여 분석하는 사례는 본 연구가 최초라 할 수 있다.

3.3 분석 데이터

대표 정형 데이터인 BC 카드의 카드거래데이터(일평균 10백만건)와 LG CNS가 크롤링하고 있는 소셜 미디어 데이터(일평균 8십만건)를 활용하여 연구를 진행하였다 [Fig. 5]. 소셜 데이터만 활용한 업무 분석 또는 카드데이터만 활용한 업무 분석은 기존에도 있었으나, 이번에는 업계 최초로 형태가 다른 두 데이터의 융합을 시도하였다.

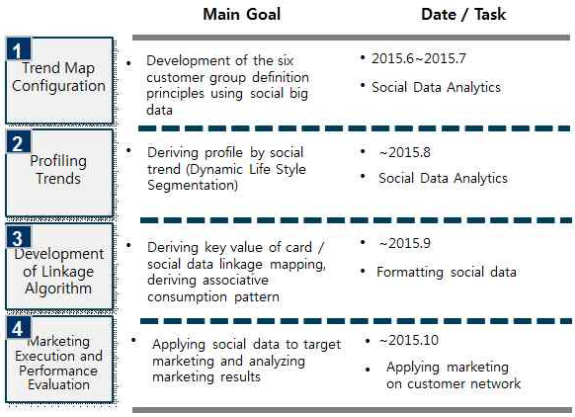


[Fig. 5] Credit Card Transaction Data and Social Media Data

3.4 분석 프로세스

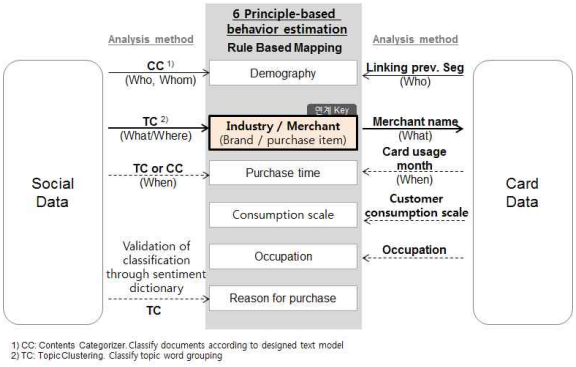
첫 번째 단계에서는 지난 5년 동안 BC카드가 추적 관리해온 소비트랜드 정보를 기반으로 트랜드 맵을 구성한다. 이 단계에서는 연관어 분석 등을 통해 트랜드별로 매핑되는 키워드를 나열하여 시각화된 맵을 구성한다. 두 번째 단계에서는 트랜드 프로파일링으로 소셜 데이터를 정형화하는 작업이다. 트랜드를 분석하는 목적은 향후 카드사 마케팅 활용에 있으므로 이의 사전 작업 격인 프로파일링 작업에서는 각 트랜드별 속성을 육하원칙에 기반 하여 행태를 정의하는 데 있다. 세 번째 단계는 카드와 소셜 데이터를 연계할 수 있는 키값을 도출하고, 통계방법을 활용하여 두가지 데이터를 연계하여 패턴을 도출한다. SNS, 블로그 등 소셜 데이터 분석을 통한 라이프스타일 기반 소비행태를 분석하고, 트랜드를 대표하는

집단의 속성을 신용카드 거래 데이터를 분석한다. 마지막으로는 빅데이터 분석으로 도출된 소셜 트랜드를 기반으로 마케팅 캠페인을 기획하고 실행하는 단계이다. 마케팅 캠페인의 실행 후에는 전년 혹은 전년대비 매출을 바탕으로 성과를 평가한다 [Fig. 6].



[Fig. 6] Analysis Process

분석과정에서는 분석에 필요한 카드 거래 데이터와 소셜 데이터를 수집, 가공하고 데이터를 정제하고 활용 가능한 정보를 탐색 및 분석하는 사전 단계를 거쳐야 한다. 분석 시에는 고객들의 라이프 스타일 변화를 토대로 하여, BC카드가 카드 거래 데이터 분석을 통해 보유하고 있는 고객 세그멘테이션 정보를 활용한다. 소셜데이터와 카드데이터의 융합을 위해서는 주성분 분석 결과를 매핑하여 매칭도가 높은 주성분 집합을 도출하기 위해, '주성분 분석'과 '유클리디안 거리' 방법을 활용한다. 주성분 분석은 많은 변수의 분산 및 공분산 구조를 간결하게 요약하여 표현하는 주성분을 선형결합을 통해 추출하는 통계기법으로 카드와 소셜데이터의 요약에 활용되었다 [21]. 유클리디안 거리는 중심 값에서 변수가 갖는 거리를 측정하는 통계 분석 기법으로 본 연구에서는 소셜데이터와 카드데이터의 주성분간 거리를 측정하여 가까울수록 소셜 데이터에서 언급하는 업종비중과 카드 데이터 업종매출 비중이 유사한 것으로 볼 수 있다. 업종 분류 그림에서 나와 있듯이, 카드사의 매출데이터와 소셜 데이터의 상품 관련 데이터를 결합하여 분석하기 위해서는 각 데이터에서 언급하는 업종을 상호 연계할 수 있어야 한다. 이를 위해 기존의 BC카드의 업종을 재분류하여 신업종 코드를 생성하였다 [Fig. 7]. 분석과 마케팅 과정을 설명하기 위해 아이들을 위해 돈을 쓰는 주체가 부모이외에도 친가와 외가의 할아버지, 할머니가 있다는 의미를 가진 "식스포켓(Six Pocket)" 관련 트랜드를 정의하여 활용한다.



[Fig. 7] Industry Classification

4. 소비트랜드 분석 모형

4.1 분석 과정 및 결과

식스포켓 트랜드와 관련 주성분 분석을 수행한 결과, 분식, 뷔페 등의 외식과 주로 관련된 업종으로 구성된 주성분, 전시 및 공연체험 업종 위주의 주성분, 유아용품 및 잡화 업종 위주로 구성된 주성분 등이 도출되었으며, 카드데이터와 소셜 데이터 각 영역에서 유사한 주성분 들이 관측되었다 [Fig. 8].

카드 데이터와 소셜데이터의 각각의 주성분에 대하여 유클리디안 거리를 측정한 결과 비교적 작아서 유사한 업종으로 구성된다고 판단되는 각 영역의 주성분을 추출한다. 어떤 유형의 업종의 비중들로 구성된 한 주성분은 하나의 소비패턴을 나타낸다고 할 수 있다. 유사한 주성분으로 판단되는 소셜데이터와 카드데이터의 소비패턴은 공통되는 여러 개의 업종을 공유할 수 밖에 없으며 [Fig. 9]에 주요한 6개의 공통되는 소비패턴을 나타내었다. 예를 들어, 카드데이터 주성분 분석결과 (패스트푸드, 전시 공연/체험, 유아용품, 의류/잡화, 자전거)를 주성분으로 하는 소비패턴과 (의류/잡화, 자전거, 학용품)을 주성분으로 하는 소셜데이터 소비패턴은 유사한 집단이라고 할 수 있으며 (의류/잡화, 자전거)를 공통으로 하는 소비패턴으로 정의할 수 있다. 카드 데이터 분석으로 알 수

없는 추가적인 관심사를 소셜 데이터에서 확인할 수 있으며 이를 마케팅에 적극적으로 활용한다.

Card data

Six-pocket new industry classification	PC1	PC2	PC3
Education_Book			
Education_ Private Education	-0.271		-0.158
Play experience		0.123	0.354
Baby goods		-0.106	0.322
Clothing _		-0.146	0.338
Bicycle		0.205	-0.539
Toy			
Exhibition_Peformance Experience		0.216	-0.406
Stationery			-0.195
Southeast Asia			0.117
Snack bar	0.259	0.296	0.264
Buffet	0.435	-0.265	
Western food	0.415	-0.211	
Japanese	0.209	-0.353	
Chinese	0.282	0.179	
Cafe _ Dessert	0.327	0.341	
Fastfood	0.263	0.547	
Korean	0.437	-0.267	-0.119

Social Data

Six-pocket new industry classification	PC1	PC2	PC3
Education_Book			
Education_ Private Education	0.111		-0.158
Play experience		0.123	0.354
Baby goods		-0.106	0.322
Clothing _		-0.146	0.338
Bicycle	0.278	0.205	-0.539
Toy			
Exhibition_Peformance Experience		0.216	-0.406
Stationery			-0.195
Southeast Asia			0.117
Snack bar	0.209	0.296	0.264
Buffet	0.418	-0.265	
Western food		-0.211	
Japanese	0.212	-0.353	
Chinese	0.156	0.179	
Cafe _ Dessert	0.327	0.341	
Fastfood	0.638	0.547	
Korean	0.234	-0.267	-0.119

[Fig. 8] Principal Component Analysis Example

[Fig. 9] Euclidean Distance Analysis Example

4.2 마케팅 캠페인 진행 및 성과

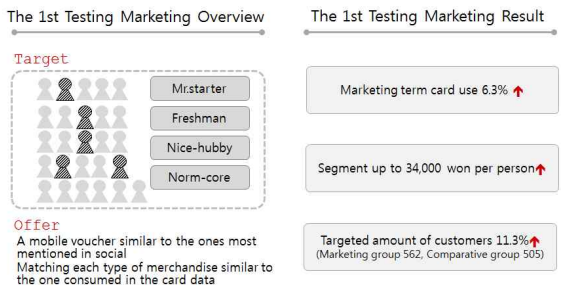
카드데이터와 소셜데이터의 융합을 기반으로 한 트랜드 도출 후에는 이를 바탕으로 마케팅 캠페인을 기획하고 실행한다. 융합 트랜드 마케팅의 성과를 다각적으로 판단할 수 있도록 매회 다른 특성을 가진 세그먼트를 대상으로 캠페인을 실시하여 객관적인 성과측정이 될 수 있도록 하였다. 캠페인은 총 3회가 실시 되었으며, BC카드의 회원사인 K 은행과, D은행에서 2015년 9월~10월 사이에 시행되었다 <Table 1>.

1차 테스트 마케팅은 기존 BC카드의 26개 고객 세그멘테이션 중 4개의 세그멘테이션 고객을 대상으로 시행되었다. 세그멘테이션은 새내기 직장인(Mr. Starter), 신입생(Freshman), 좋은 남편(Nice-bubby), 노멀 코어(Norm-core), 식스포켓(Six-pocket)이다. 노멀코어는 노멀(Normal)과 하드코어(Hardcore)의 합성어로 평범함을 표방하지만 평범하지 않은 패션을 추구하는 세그멘테이션을 의미한다. 싯스포켓은 조카를 두었거나 아이들에게 선물을 할 만한 동기가 있는 연령 혹은 이러한 소비취향을 가진 것으로 판단되는 세그멘테이션이다.

<Table 1> Marketing Campaign List

	1st marketing	2nd marketing	3rd marketing
Instituion	K bank	K bank	D bank
Duration	'15.9.12~9.30	'15.10.1~10.11	'15.10.20~10.31
Target customer	10,000 people in the high trendy propensity segments	Three segments with card/social linked consumption patterns 20,000 people	Five segments with card/social linked consumption patterns 70,000 people

마케팅은 소셜데이터에서 많이 언급된 업종과 유사한 소비패턴을 갖는 카드 소비 업종 패턴을 매핑하여 카드 고객 약 1만명을 대상으로 목표금액(10~30만원)을 사용하면 할인쿠폰을 사용할 수 있도록 하는 방법으로 진행하였다. 이러한 마케팅을 시행하는 기간동안 카드이용금액은 전년 대비 6.3%, 인당 이용금액은 최대 3만 4천원이 증가하였으며, 마케팅 목표 금액 달성 고객수는 전년도 캠페인 대비 11.3%가 증가하는 성과를 거두었다 [Fig. 10].



[Fig. 10] Marketing Campaign Result(1st)

2차 테스트 마케팅은 약 2만명을 대상으로 기존 카드 세그멘테이션과의 연계없이 이번 과제의 분석 결과만을 활용하여 카드/소셜 연관 업종 소비 패턴을 보이는 집단을 타겟으로 하고 소셜 상에서 관심있게 언급한 업종, 카드데이터에서 소비를 보인 업종을 각각 마케팅 오퍼로 매칭하여 마케팅을 10.1~11일(총 12일)간 시행하였다. 동일한 소비성향을 보이는 고객 세그먼트를 마케팅을 시행한 실험집단과 시행하지 않은 통제집단으로 나누어 캠페인의 성과를 측정해 본 결과, 실험집단은 카드 이용금액이 약 9억이 증가하였으며, 인당 이용금액은 약 4만원이 증가하였음을 확인하였다. 또한 마케팅 목표에 도달한 고객이 통제집단대비 3백명이상 많아 캠페인이 성과가 있음을 객관적으로 증명하였다. 3차 테스트 마케팅은

D은행 고객을 대상으로 시행되었다. 약 7만명을 대상으로 진행하였는데 12일간 진행한 단기 마케팅임에도 불구하고 오퍼 반응율 평균 6.3% 달성하여 약 7천 6백만원의 수익을 창출하였다 <Table 2>.

<Table 2> Offer Response

그 룹	Offer response				Total sales (B)	Sales profit (C=B*1%)	Revenue C-A)
	Target (per.)	Response (per.)	Response rate	Offer cost(A)			
❶Clothing stuff, Bicycle	5,707	383	6.7%	3,830,000	1,100,076,420	11,000,764	7,170,764
❷Clothing stuff, Baby goods	2,437	182	7.5%	1,820,000	466,100,010	4,661,000	2,841,000
❸Education book, Play experience	7,378	462	6.3%	4,620,000	1,331,379,255	13,313,793	8,693,793
❹Play, Exhibition performance experience	30,001	2,058	6.9%	20,580,000	5,426,628,118	54,266,281	33,686,281
❺Baby goods, Exhibition performance experience	25,536	1,405	5.5%	14,050,000	3,860,221,125	38,602,211	24,552,211
Sum	71,059	4,490	6.3%	44,900,000	12,184,404,928	121,844,049	76,944,049

총 3회에 걸쳐 진행한 마케팅 캠페인의 성과를 확인하기 위해 전월, 전년동기 대비 기존 정기 캠페인과 비교한다. 비교는 마케팅 결과를 분석할 때 가장 일반적으로 사용하는 마케팅 승수를 계산하여 비교하였다. 마케팅 승수는 마케팅에 투여한 비용 1원당 얼마만큼의 매출증대 효과가 있었는지를 측정하는 지표로, 마케팅기간에 발생한 카드 이용금액에서 전월 동기간 발생한 카드이용금액을 뺀 차액을 마케팅에 투여한 비용으로 나누어 계산한다. 전반적인 마케팅 승수는 42~101로 캠페인의 성과가 있음을 확인하였고 상세결과는 <Table 3>에 제시하였다.

<Table 3> Final Campaign Result

	Campaign duration (Billion Won)	Previous month (Billion Won)	Previous year (the same month, Billion Won)	Marketing multiplier		Marketing cost (Million won)
				Last Month	Previous year	
1st marketing	9.8	7.9	7	**33.90**	49.10	5.6
2nd marketing	59.1	55.5	38.9	**17.81**	101.66	1.9
3rd marketing	127.2	113.8	108	**29.67**	42.66	4.4

- Previous month's marketing multiplier = (Campaign period amount-Previous month amount)/Marketing cost
- Existing six regular upsell marketing Average marketing multiplier per month: **15.86**
(Campaign code: average of 6 campaigns including 4 performance-enabled campaigns, 2 recruitment card campaigns)

5. 결론

본 연구에서는 기존에는 시도하지 않았던 정형, 비정형의 형태를 가진 카드와 소셜 데이터를 '업종'을 매개로 하여 결합하고 새로운 트랜드를 도출한 후 실제 카드사의 마케팅에 적용하여 지속적으로 매출에 긍정적인 영향이 있음을 확인하는 성과를 거두었다. 식스포켓으로 상징되는 키즈마켓에 대한 카드와 소셜 데이터를 분석하여 서로 유사한 매출 패턴을 가지는 상품군(의류잡화와 자전거 등)을 도출하고, 이를 마케팅 오퍼로 응용하여 관련 소비를 유도한 결과 인당 이용금액이 상승하는 성과를 거두었다. 카드와 소셜 연관업종에 대한 동일한 소비 패턴을 가지는 집단으로 한 마케팅에서도 인당 이용금액이 상승하는 성과를 보여 카드-소셜 데이터 연계의 성과를 증명하였으며, 최종적으로 카드와 소셜데이터에서 도출된 연관소비 패턴을 반영한 마케팅에서도 오퍼 반응율 6.3%를 달성하여 데이터 융합으로 인한 가능성을 체계적으로 증명하였다.

카드사의 정형데이터는 정확한 고객개인과 소비정보를 포함하고 있어 분석이 용이하지만 어떠한 상황에서 어떤 행태의 소비를 하는지 소비양태를 파악하기 어려울 뿐더러 결과론적인 후행데이터여서 현재의 소비트랜드와 이에 적합한 마케팅을 효과적으로 하기 힘든 면이 있다. 반면, 소셜데이터는 사회문화적인 이슈와 결부되어 변화하는 현 시점의 소비트랜드를 실시간으로 반영하고 있어 마케팅 측면에서 유리한 면이 있지만, 정확한 화자 주체를 파악하기 어려워 분석과 마케팅에 활용하기 어려운 측면이 있었다.

본 연구는 비정형 소셜데이터를 BC카드가 보유한 거래 데이터와 융합하여 이를 실제 마케팅에 적용한 금융업 최초의 사례로 빅데이터의 새로운 가능성을 보여 주었다. 소비자에 관한 선행데이터인 소셜데이터와 후행데이터인 카드데이터를 융합하여 소비자에 대한 상호보완적인 분석과 선행적인 마케팅을 할 수 있도록 방법론을 제시하였다. 소비자의 구매상품과 이것이 속한 업종을 매개체로 하여 융합한 양쪽의 프로파일은 현재 소비자의 관심을 최소 한 달에서 수 개월 전에 예상할 수 있게하고 이 소비트랜드에 관해 효과적으로 반응할 마케팅 집단을 선정할 수 있도록 한다. 일반적으로 소셜데이터는 양은 많지만 불명확하여 금융회사에서는 분석용으로 기피되고 있다. 그러나 대표적인 비정형 빅데이터인 소셜데이터도 정교한 분석과 융합방법론이 뒷받침된다면 유용하게 활용될 수 있다는 것을 본 연구결과는 보여주고 있다.

또한 본 연구는 담당자의 역량에 의존하던 기존 마케

팅의 한계를 소셜데이터와 카드 매출데이터를 결합한 인사이트를 통해 극복한 의의가 있다. 이러한 점에서 이번 연구는 소셜 데이터의 트랜드 분석을 통해 빅데이터 활용법을 제시하고 매출신장이라는 실질적인 성공을 거둠으로 인해 빅데이터 성공사례를 창출한 점은 보수적인 금융업의 혁신을 유도하고 있다는 점에서 시사점이 크다. 사회의 소비 트랜드를 반영하고 있는 소셜 데이터를 활용한 정교한 마케팅 성공 사례라는 점과 개인정보보호법 등 각종 규제를 준수하면서도 이종 데이터의 융합 등 어려움을 극복할 수 있는 방안을 제시한 점도 본 연구의 공헌이다. 연구측면에서도 주성분 분석과 유클리디안 거리를 통해 정형데이터와 비정형데이터를 연계 분석할 수 있는 방법론을 제공한 점과, 카드업계에서 활용되는 마케팅 승수를 통한 성과비교를 제시한 점도 성과로 꼽을 수 있다. 본 연구는 카드업종 뿐만 아니라 다른 업종에서도 자사데이터와 소셜 데이터를 연계 분석하기 위한 레퍼런스를 제공한 의의가 있다. 한편으로 본 연구에서 기존 캠페인 방식과의 성과 비교 같은 좀 더 정교한 방법을 통한 성과증명이 미흡한 점은 한계로 남는다.

카드사의 매출데이터는 소셜 데이터 뿐만 아니라, 공공데이터 및 각 기관과의 제휴를 통해 이종 데이터의 융합으로 의미있는 인사이트를 도출하거나 이를 지수화하여 지속가능한 형태로 제공하는데 활용할 수 있다. 카드데이터를 비식별 조치하여 고객의 세그먼트 단위로 소비트랜드를 분석하여 유통기업 등에 제공한다면 이들 기업의 효과적인 마케팅에도 활용될 수 있다. 또한 카드데이터는 실물경제 동향 분석에 유용하게 활용될 수 있으며, 증권사가 보유한 증권 거래데이터와 융합할 경우 경제현황에 대한 심층적인 분석과 예측도 가능해진다. 앞으로 카드데이터를 다양하게 활용하여 타 산업에 제공가능한 비즈니스 모델을 제시한다면, 의미 있는 마케팅 컨설팅 사업과 유료 비즈니스 모델의 창출로 융합 신산업 활성화에도 큰 공헌이 될 것이다.

ACKNOWLEDGMENTS

This paper was supported by Wonkwang University in 2017.

REFERENCES

[1] D. Laney, “3D Data Management: Controlling Data Volume, Velocity and Variety”, Gartner, 2001.

[2] NRI, “The Advent of Big Data Era”, Nomura Research Institute, 2012.

[3] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, “Big Data: The Next Frontier for Innovation, Competition, and Productivity”, McKinsey Global Institute, 2011.

[4] MSIP, "Big Data Industry Development Strategy", The Republic of Korea Government, 2013.

[5] N. Ashish, V. Dan, "Worldwide Big Data Technology and Services Forecast 2015-2019", IDC, 2015.

[6] NIA, 2016 Korean Big Data Market Review, National Information Society Agency, 2017.

[7] J. S. Hwang, S. H. Kim, and H. Lee, "Breaking the Myths of the IT Productivity Paradox", KSII Transactions on Internet and Information Systems, Vol. 9, No. 1, pp.466-482, 2015.

[8] K. I. Kim, "The Impact of several management tools and techniques adoption on strong small business enterprises'Performance", Journal of IT Convergence Society for SMB, Vol. 6, No. 3, pp.7-12, 2016.

[9] K. I. Kim, "A Model of the influence of IMS operation on organizational performance", Journal of IT Convergence Society for SMB, Vol. 6, No. 1, pp.1-5, 2016.

[10] S. H. Kim, S. Y. Jang, and G. O. Jang, "Information Technology Adoption and Success Research Review in Small and Medium-sized Enterprises", Journal of Digital Convergence, Vol. 11, No. 3, pp.199-207, 2013.

[11] S. H. Kim, and S. Y. Jang, "Determinants of Software-as-a-Service Adoption Intention Among Small and Medium-sized Korean Firms", Korean Management Science Review, Vol. 30, No. 2, pp.11-30, 2013.

[12] Y. S. Jeong, "U-healthcare Service Management Scheme for Big Data of Patient Information", Journal of IT Convergence Society for SMB, Vol. 5,

No. 1, pp.1-6, 2015.

[13] K. S. Shin, H.J. Park, S. H. Kim, Big Data Adoption Strategy in Finantial Industry, National Information Society Agency, 2016.

[14] D. Turner, M. Schroeck, and R. Shockley, Analytics: The Real-World Use of Big Data in Financial Services, IBM® Institute for Business Value, 2016.

[15] NIA, 2015 Korean Big Data Market Review, National Information Society Agency, 2016.

[16] S. H. Lee, and D. W. Lee, "FinTech - Conversions of Finance Industry based on ICT", Journal of the Korea Convergence Society, Vol. 6, No. 3, pp.97-102, 2015.

[17] R. Baik, and J. H. Ryu, "An Analysis of Convergence Relation on Economic Activity with Credit cards in Korea and China & A Development of the Algorithm on economic trend Estimation", Journal of the Korea Convergence Society, Vol. 7, No. 4, pp.9-17, 2016.

[18] K.G. Min, http://www.econovill.com/news/articleView.html?idxno=298758, Economic Review Article, 2016.

[19] E. J. Noh, http://www.focus.kr/view.php?key=2016113000161540599, Focus Nesws Article, 2016

[20] H. Chen, R. H. L. Chiang, and V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," MIS Quartly. Vol. 36, No. 4, pp. 1165-1188, 2012.

[21] H. Baars, and H. G. Kemper, "Management Support with Structured and Unstructured Data-an Integrated Business Intelligence Framework," Information Systems Management, Vol. 25, No. 2, pp. 132-148, 2008.

[22] A. Syed, K. Gillela, and C. Venugopal, "The Future Revolution on Big Data," . International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, No. 6, pp. 2446-2451, 2013.

[23] A. Gandomi, and M. Haider, "Beyond the Hype: Big Data Concepts, Methods, and Analytics," International Journal of Information Managemen, Vol. 35, No. 2, pp. 137-144, 2015.

[24] K. S. Shin, S. M. Chai, H. J. Park, N.O. Jo, S. A. Shin, and S. H. Kim, "Development of a Big Data Capability Assessment Model", Journal of Information Technology and Architecture, vol.13, no.2, pp. 271-280, 2016.

[25] S. W. Lee and S. H. Kim, "Finding Industries for Big Data Usage on the Basis of AHP", Journal of Digital Convergence, Vol. 14, No. 7, pp. 21-27. 2016.

[26] S. H. Kim, S. B. Park, and Y. G. Lee, "A Development of a Evaluation Framework for Public Sector ICT Adoption: Focused on Big Data, Cloud, Internet of Things, Journal of Information Technology and Architecture, Vol. 12, No. 3, pp. 419-428, 2015.

[27] Y. Jung, M. Suk, C. Kim, "A study on the success factors of Bigdata through an analysis of introduction effect of Bigdata", Journal of Digital Convergence, Vol 12, No. 11, pp. 241-248, 2014.

[28] K. S. Noh, "Educational Policy Proposals through Analysis of the Perception of Bigdata for University Students", Journal of Digital Convergence, Vol. 13, No. 11, pp.25-33, 2015.

[29] Y. M. Kim, "Changes and Strategies of the Government Service Paradigm through Using Big Data: Focused on Disaster Safety Management in Seoul City", Journal of Digital Convergence, Vol. 15, No. 2, pp.59-65, 2016.

[30] M. J. Choi, and K. S. Noh, "Exploratory Study on Crime Prevention based on Bigdata Convergence - Through Case Studies of Seongnam City -", Journal of Digital Convergence, Vol. 14, No. 11, pp.125-133 2016.

[31] K. S. Noh, "A Study on Utilization Strategy of Bigdata for Local Administration by Analyzing Cases", Journal of Digital Convergence 12(1), 2014.

[32] K. S. Noh, S. Park, "An Exploratory Study on Application Plan of Bigdata to Manufacturing Execution System", Journal of Digital Convergence, Vol. 12, No. 1, 2014.

[33] J. S. Kim, and W. S. Cho, "Data analysis of 4M data in small and medium enterprises", Journal of the Korean Data & Information Science Society,

Vol. 26, No. 5, pp.1117 - 1128, 2015.

[34] J. W. Gu, J. H. Lee, M. S. Chung, and J. Y. Lee, "Electric Vehicle Technology Trends Forecast Research Using the Paper and Patent Data", Journal of Digital Convergence, Vol. 15, No. 2, pp.165-172, 2017.

[35] H. W. Park, and K. H. Choi, "Doing Social Big Data Analytics: A Reflection on Research Question, Data format, and Statistical Test-Convergent Aspects", Journal of Digital Convergence, Vol. 14, No. 12, pp.591-597, 2016.

[36] S. D. Yoo, K. D. Choi, and S. Y. Shin, "Characterizing Business Strategy in a New Ecosystem of Big Data", Journal of Digital Convergence, Vol. 12, No. 4, pp.1-9, 2014.

[37] S. H. Namn, and K. S. Noh, "A Study on the Effective Approaches to Big Data Planning", Journal of Digital Convergence, Vol. 13, No. 1, pp.227-235, 2015.

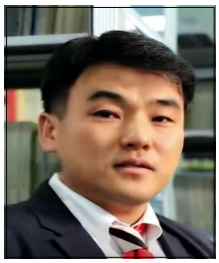
김 성 현(Kim, Sunghyun)

- 2005년 2월 : 고려대학교 경영학과 (경영학석사)
- 2013년 2월 : 성균관대학교 경영학과 (경영학박사)
- 2000년 1월 ~ 2005년 11월 : 삼성 SDS IT컨설턴트
- 2005년 11월 ~ 현재 : 한국정보화진흥원 수석연구원
- 관심분야 : 정보기술 도입, 정보화성과관리, 빅데이터
- E-Mail : kimcon@nia.or.kr

장 석 호(Chang, Sokho)

- 1998년 2월 : 강원대학교 전자공학과 (공학박사)
- 2005년 2월 : 서울대학교 경영대학원 (경영학석사)
- 2012년 3월 ~ 현재 : 비씨카드 빅데이터센터 센터장
- 관심분야 : 마케팅, 빅데이터, 고객관계관리
- E-Mail : sokhochang@bccard.com

이 상 원(Lee, Sangwon)



- 2002년 2월 : 한국과학기술원 경영정보학과 (경영학석사)
- 2009년 8월 : 한국과학기술원 경영공학과 (공학박사)
- 2011년 3월 ~ 현재 : 원광대학교 컴퓨터소프트웨어공학과 부교수
- 관심분야 : 데이터공학, 정보공학, 지식공학
- E-Mail : sangwonlee@wku.ac.kr