

텍스트 마이닝을 이용한 이익조정 연구동향 토픽모델링

김지연¹, 나홍석^{2*}, 박경환³

¹고려사이버대학교 융합정보대학원 석사, ²고려사이버대학교 융합정보대학원 교수, ³고려사이버대학교 경영학과 교수

Topic Modeling of Profit Adjustment Research Trend in Korean Accounting

JiYeon Kim¹, HongSeok Na^{2*}, Kyung Hwan Park³

¹Graduate Student, Graduate School of Interdisciplinary Information Studies, The Cyber University of Korea

²Professor, Graduate School of Interdisciplinary Information Studies, The Cyber University of Korea

³Professor, Dept. of Business Administration, The Cyber University of Korea

요약 본 연구는 이익조정을 연구 주제로 국내 회계학 관련 논문들의 세부 연구 동향 파악을 목적으로 한국학술지인용 색인(KCI)의 회계학 또는 경영학에 속하는 학술지에 게재된 논문 초록을 텍스트 마이닝 기법을 이용하여 분석하였다. 분석 결과, 이익조정 관련 회계학 논문들이 감사 및 감사보고서, 법인세 및 부채비율, 기업의 전반적인 경영전략, 재무제표와 회계기준의 4가지 연구 영역으로 나누어지는 것을 확인하였으며, 재무제표와 회계기준이 주된 토픽일 것이라는 예측과는 달리 감사가 가장 많이 연구된 토픽으로 분석되었다. 이어 토픽별 논문 수를 기준으로 토픽 트렌드를 확인하고 특이사항에 대한 원인을 분석하였으며, K-IFRS의 도입이 이익조정 연구에 미친 영향을 확인하였다. 본 연구는 이익조정 에 대한 자세한 연구 동향에 대한 정보를 제공하고 회계학 문헌분석 연구방법으로 텍스트 마이닝 기법을 제시하였다. 또한 정책결정자 및 기업 실무 담당자가 이익조정 관련하여 회계기준 외에 감사 등 4개 토픽별로 추가적으로 고려하여야 할 사항과 그 추세를 파악할 수 있도록 하였다.

주제어 : 빅데이터, 텍스트마이닝, 토픽모델링, 연구동향, 회계학, 이익조정

Abstract This study identifies the trend of Korean accounting researches on profit adjustment. We analyzed the abstract of accounting research articles published in Korean Citation Index (KCI) by using text mining technique. Among papers whose themes were profit adjustment, topics were divided into 4 parts: (i) Auditing and audit reports, (ii) corporate taxes and debt ratios, (iii) general management strategy of companies, and (iv) financial statements and accounting principles. Unlike the prediction that financial statements and accounting principles would be the main topic, auditing was analyzed as the most studied area. We analyzed topic trends based on the number of papers by topic, and could figure out the impact of K-IFRS introduction on profit adjustment research. By using Big Data method, this study enabled the division of research themes that have not been available in the past studies. This study enables the policy makers and business managers to learn about additional considerations in addition to accounting principles related to profit adjustment.

Key Words : Big data, Text mining, Topic modeling, Research trend, Accounting, Profit adjustment

*Corresponding Author : HongSeok Na(hsna99@cuk.edu)

Received November 26, 2020

Accepted January 20, 2021

Revised January 4, 2021

Published January 28, 2021

1. 서론

본 연구에서는 텍스트 마이닝 분야의 LDA 토픽모델링(Topic Modeling) 기법을 적용하여 '이익조정'이라는 회계학의 특정 연구주제를 세부적으로 분석하고, 해당 연구주제 내에서의 연도별 토픽 트렌드를 분석하였다.

연구주제는 회계학이라는 학문의 연구 대상이 무엇인가, 그리고 회계학이 다른 학문과 차별화되는 점이 무엇인가에 대한 답을 제공한다. 국내 회계학 논문들은 연구주제의 편중 현상이 두드러지게 나타나는데[1], 그 중에서 '이익조정' 관련 연구는 지속적으로 높은 관심을 가져온 주제이다. 국내 회계학 연구에 대해 40여년의 기간을 대상으로 광범위하게 분석한 최원석 등(2015)의 연구에 의하면, '이익조정'은 재무회계 중 1위(66건, 20.50%), 세무회계 중 2위(7건, 8.4%), 회계감사 중 2위(6건, 8.2%)로 회계학의 다양한 분야에 걸쳐 높은 빈도로 연구되었다[2]. 권선국과 조성표(1997), 권수영과 정종욱(1999)의 연구에서도 공통적으로 이익조정이 회계학의 주요 주제로 연구되고 있음을 분석한 바 있다[3,4]. 회계기준 위반으로 규제당국에 의해 적발될 경우에는 기업에 막대한 소송비용과 추가적인 감시비용을 야기할 뿐만 아니라 규제기관의 지속적인 감시로 인해 기업 활동이 제약을 받을 수 있다는 점에서[5], 이익조정은 정책적, 실무적인 관점에서 중요한 주제이다.

그러나 선행 연구 동향을 살펴보면, 이익조정은 국내 회계학 분야에서 중요한 연구주제임에도 불구하고 이에 초점을 맞추어 독립적으로 분석되지 않았다. 특히, 이익조정이 어떠한 토픽으로 세부적으로 나누어지고 있으며 그 흐름이 어떠한지에 대해 파악하는 시도는 아직까지 이루어지지 않았다. 한중수와 송민섭(2018)의 견해에 따르면, 이익조정은 매우 중요한 회계적 사안임에 누구나 동의하나 많은 논문들이 '경영자가 기회주의적 유인에 의하여 이익조정을 한다'는 사실 이외의 어떠한 추가적인 사실을 밝혔는지는 의문시된다고 하였다[1].

경영자는 자신에게 허용된 회계 재량권을 이용하여 보고이익 수치를 필요에 따라 조정하는 경향이 있는 것으로 알려져 있다[5]. 오래 전부터 회계변경을 이용한 이익조정에 대한 연구가 있었으며[6], 주로 회계이익에 포함된 이익조정액의 측정에 초점을 맞추어 연구가 진행되어 왔다[7]. 그러나 이익조정은 회계분야에서 매우 광범위하게 연구가 진행되어 온 주제이기 때문에 단순히 '재무제표나 회계기준에 한정'되어 연구가 진행되어 왔다고 단정할 수는 없을 것이다. 이익조정은 회계기준 뿐만 아니라

감사품질, 경영정책, 법인세 효과 등 여러 측면에서 복합적인 요소가 결합되어 있을 것으로 판단된다. 그럼에도 아직까지 회계학 연구에서 '이익조정'을 키워드로 어떻게 토픽들이 구성되어 있는지 밝히는 연구논문은 찾아볼 수 없다. 이에 본 연구는 그동안의 회계학 분야에서 이루어져 온 이익조정에 대한 연구들을 심층탐색을 통해 종합적으로 새롭게 바라볼 수 있게 한다는 점에서 연구의 의미가 있다.

한편, 회계학 연구에서 '이익조정'에 대한 관련 토픽 분석이 종합적으로 수행되지 못한 이유는 연구자가 수작업으로 진행되는 문헌분석 방법의 한계 때문인 것으로 판단된다. 우리나라 기존의 회계학 분야 연구 동향 분석은 수작업을 통한 문헌분석 방식이 주를 이룬다. 먼저 회계학 분야의 대표적인 학술지 1~2개를 선택하고, 일정 기간 동안 특정 회계학 학술지에 발표된 논문들을 대상으로 하여 이들의 연구주제를 분석하는 방식이다. 이 때 1차적으로 재무회계나 세무회계와 같이 넓은 연구분야로 구분하고, 2차적으로 각 연구분야에서 주로 어떠한 주제가 연구되었는지 분석하였다. 이익조정, 정보내용 등은 이렇게 분석된 연구주제에 해당한다. 하지만 여기에서 한 단계 더 나아가 이익조정 등 특정 연구주제 내에서의 연구 흐름에 대해서는 기존의 문헌분석 방법으로는 한계가 있다.

회계학의 문헌분석 연구에서 한계점으로 자주 언급되는 부분은 첫째, 일부 학술지만을 분석대상으로 하였다는 점[3,8,9], 둘째, 분석 대상 기간을 특정하였다는 점[8,9], 셋째, 자료의 수집과 분석을 수작업으로 하였다는 점[10], 넷째, 위의 한계점들과 관련하여 분석대상 포함 여부에 연구자의 주관적인 판단이 개입된다는 점이다[3,8,9].

본 연구에서는 기존의 회계학 분야 문헌분석 연구에서 한계점으로 언급되었던 부분을 보완하는 방법으로 빅데이터 분석 기법을 사용한다. 빅데이터 분석은 방대한 양의 데이터를 객관적으로 수집, 분석하여 여러 가지 사용 방식에 따라 새로운 가치를 창출할 수 있는 도구이다[11]. 특히 빅데이터 분석 기법 중 하나인 텍스트 마이닝(Text-mining)은 인간의 언어로 쓰인 비정형 텍스트에서 자연어 처리기술을 이용하여 유용한 정보를 추출하는 것으로 비정형 텍스트의 연계성을 파악하여 분류 혹은 군집화하거나 요약하는 등 빅데이터 속에 숨겨진 의미 있는 정보를 발견하기 위해 주로 사용된다[12]. 텍스트 마이닝 기법을 활용한 연구 동향 분석은 기존의 연구자 중심의 문헌분석의 한계와 연구자의 주관이 반영되는 단

점을 극복할 수 있다[11].

또한 텍스트 마이닝 기법을 활용하는 경우 이익조정 관련 토픽들이 시간의 흐름에 따라 어떻게 변동하는지 경향을 파악할 수 있다. 즉, 시대별로 어떤 토픽이 주된 연구 주제인지를 확인할 수 있다. 2011년 한국채택국제 회계기준(K-IFRS) 도입으로 회계정책 및 회계처리방법의 선택권이 경영자에게 상당 부분 부여되었는데[13], 이에 따라 K-IFRS 도입 후 회계학 연구에서 이익조정 관련 토픽들의 변화가 시간의 흐름에 따라 어떻게 진행되었는지 심층적으로 분석 가능하다.

본 연구는 ‘이익조정’이라는 주제에 대해 기존 문헌분석에 의한 연구동향 분석에서 한계로 지적된 점을 텍스트 마이닝 기법을 도입하여 해결하고자 한다. 회계학 분야의 ‘이익조정’ 관련 토픽들을 구조화 하여 분석하고, 시간의 흐름에 따라 토픽들이 어떻게 변화되어 왔는지 종합적으로 연구하여 그 결과를 제시하였다는 점에서 선행 연구와 차별되는 연구의 의의가 있다.

본 연구는 1장 서론에 이어 2장에서는 선행연구를 살펴본다. 3장은 토픽모델링 기법을 설명하고, 4장에서는 이 기법을 사용하여 이익조정의 토픽별 키워드와 연도별 토픽 트렌드를 분석하였다. 마지막으로 결론에서 연구결과를 요약하고 시사점과 기대효과를 제안하였다.

2. 선행연구 검토

2.1 회계학분야 이익조정 선행연구

이익조정(earnings management)이란 좁은 의미에서는 경영자의 재량권에 따라 법과 회계기준의 테두리 안에서 수행되는 정당한 행위를 의미하지만[14], 넓은 의미에서는 기업회계기준을 어겨가면서까지 수행되는 분식 회계를 통칭하는 용어로 사용된다[15]. 이익조정은 경영자가 재량적 회계선택을 통해 이익을 상향하여 보고함으로써 기업의 신용등급을 높이거나[14] 적자보고를 회피하는 수단 등으로 사용되며[15], 일반적으로 경영자의 기회주의적 유인에 의한 도덕적 대리인 문제(agency problem)로부터 발생하는 것으로 알려져 있다[1].

이익조정은 1970년대 들어 정성적 연구 대신 실증연구가 주요 회계학 연구방법이 된 이후 수십 년 동안 회계 연구의 주요 주제로 다루어져 왔는데[1], 이는 회계학 분야의 연구동향에 대한 문헌분석을 수행한 권선국과 조성표(1997), 서은경과 한인구(1998), 권수영과 정종욱(1999), 최원석 등(2015)의 연구에서도 확인된다. 그동

안의 회계학 분야의 연구동향에 대한 분석은 대부분 재무회계로의 집중화 현상을 보이며, 수십 년 동안 회계 연구의 주요 주제가 크게 변화하지 않았다는 특징이 있다 [1]. 그리고 이익조정, 가치관련성, 자본비용 등 특정 주제에 연구가 집중되어 다양한 회계적 사안이 논의되고 있지 못하다는 단점이 지적된다[1].

권선국과 조성표(1997)는 한국회계학회의 학술지인 『회계학연구』에 1977년부터 1996년까지 20년 동안 게재된 225편의 논문들을 대상으로 연구분야 및 연구 분야별 주제와 연구방법에 대하여 고찰하였다[3]. 재무회계가 54%로 회계학 연구의 주된 연구 분야이며, 이익조정은 재무회계 분야에서 주요 연구주제인 것으로 분석되었다.

서은경과 한인구(1998)는 1990년부터 1996년까지 한국회계학회의 학술지인 『회계학연구』 및 『회계저널』에 수록된 모든 논문과 한국경영학회의 학술지인 『경영학연구』에 실린 회계학 주제의 논문을 분석하였다[10]. 214편의 논문을 대상으로 저자유형, 저자 소속기관 유형, 연구주제, 연구방법론 분석 등을 수행하였다. 분석 결과 재무회계, 관리회계 및 회계감사 분야가 주된 연구 영역인 것으로 나타났으며, 재무회계의 연구 비중이 52%로 가장 많은 것으로 검토되었다.

권수영과 정종욱(1999)은 『회계학연구』, 『회계저널』, 『경영학연구』의 창간호부터 1998년 12월까지 발표된 424편의 논문을 대상으로 하여 시대별, 학술지별로 연구분야와 연구방법론을 검토하였다[4]. 분석 결과 재무회계가 분석대상 논문의 51.89%로 가장 많은 비중의 논문이 게재된 것으로 나타났으며, 이익조정은 재무회계 분야에서 주요 연구 주제인 것으로 분석되었다.

최원석 등(2015)은 한국회계학회가 발간하는 『회계학연구』와 『회계저널』의 창간호부터 2013년 12월까지 약 40여년간 게재된 논문을 대상으로 하여 내용분석, 저자 분석, 심사기간 및 인용분석 등을 수행하였다[2]. 『회계학연구』에서는 재무회계, 관리회계, 세무회계, 회계감사의 순서로 연구의 비중이 높은 것으로 관찰되었으며, 이익조정이 재무회계 분야에서 많이 연구된 주제로 확인되었다. 『회계저널』에서는 재무회계, 관리회계, 회계감사, 세무회계의 순서로 연구의 비중이 높은 것으로 관찰되었으며, 역시 이익조정과 IFRS가 재무회계 분야에서 많이 연구된 주제로 확인되었다. 이러한 연구물들의 연구방법론은 대부분 실증연구가 차지하고 있다고 보고하였다.

이처럼 회계학 분야 연구동향에 대한 앞선 문헌연구들은 1977년부터 최근 2015년까지 중 특정 기간에 진행된 연구 자료를 분석하였고, 그 숫자는 손에 꼽을 정도로 매

우 적다고 볼 수 있다. 또한 적용된 연구방법론은 수작업을 통한 문헌연구 방식이라는 점에서 체계적이지 않은 방법들이 동원되었다는 것을 알 수 있다.

이익조정을 주요 주제로 다룬 개별 연구 논문들을 살펴보면, 대부분 이익조정 동기과 그로 인한 이익조정 현상의 존재 여부를 실증하는데 집중해 온 경향이 있다[5]. 국내의 많은 연구들은 이익조정 동기가 존재하는 기업에서 경영자의 재량 및 의도에 따라서 기업의 회계이익이 조정되고 있음을 검증하고 있다[16]. 이러한 이익조정은 감사품질이 높은 감사인으로부터 외부감사를 받을 경우 유의적으로 낮게 나타나는 것으로 보고되어, 높은 감사품질이 경영자의 기회주의적 이익조정 행위를 제한한다는 연구결과도 있다[17]. K-IFRS도 이익조정 연구에서 종종 등장하는 주제이다. 상장회사 및 비상장 금융회사는 2011년부터 K-IFRS를 의무적으로 적용하는데, 이로 인해 경영 및 회계환경에 적지 않은 변화를 수반하였기 때문이다[13]. 또한 주관적 판단의 개입 여지가 큰 K-IFRS 하에서 경영자들은 명확하지 않은 회계기준을 통해 이익조정을 시도하려는 경향이 나타나기 때문에, 경영자가 주어진 재량권을 기회주의적으로 이용하여 기업의 회계이익을 조정한다는 기존의 연구 결과와 연결하여 논의되었다[16,18].

이익조정은 회계학의 중요한 연구주제이나[1], 이처럼 이익조정을 주제로 하는 개별 연구들은 이익조정 동기와 현상의 관계, 감사와 이익조정 관계, 회계적 사건과 이익조정 관계와 같이 단편적인 독립적 실증연구에 한정하여 분석이 수행되었다는 한계가 있다. 그러므로 이 주제에 대한 종합적인 고찰을 가능하게 하는 연구의 필요성이 제기된다고 볼 수 있다.

이러한 관점에서 ‘이익조정’을 주제로 하여 관련 토픽들을 탐색하여 구조화 하고 시대의 흐름에 따라 그러한 토픽들이 어떻게 변화되어 왔는지를 일목요연하게 정리할 필요가 있다. 또한 회계학에 큰 영향을 미친 것으로 알려진 2011년 K-IFRS 도입이 실제로 이익조정 관련 연구에 어떠한 변화를 가져왔는지 탐색할 필요가 있다.

본 연구는 이를 위해 회계학의 문헌연구 분석에 텍스트 마이닝을 기반으로 한 빅데이터 분석 기법을 도입하였다. 다음에서는 토픽모델링 분석 기법에 대해 알아보고, 이 기법을 통한 분석 절차와 분석 결과를 제시한다.

2.2 토픽모델링을 활용한 연구 동향 분석

토픽모델링은 데이터 마이닝의 기법 중 하나로서 비구조화된 텍스트 자료들의 뭉치로부터 의미있는 주제(토픽)

들을 추출해주는 확률모델 알고리즘이다[19]. 초기에는 확률적 잠재의미분석(pLSA) 기법이 사용되었으나, Blei가 잠재 주제의 확률적 분포에 대한 잠재의미분석의 한계점을 보완한 모델로 LDA(Latent Dirichlet Allocation) 알고리즘을 발표한 이후로는 주로 LDA 기법이 사용되고 있다[19]. 토픽에 따라 문서 집합을 재구성할 수 있으며, 유전(gene) 데이터, 이미지 및 소셜 네트워크에서 패턴을 찾는 데에도 사용된다[20]. 특히 인터넷 리뷰, 신문기사, 논문, 특허 데이터 분석에 있어 동향과약과 잠재 주제를 파악하는데 유용하게 사용되며, 학계에서도 축적된 연구 자료를 바탕으로 토픽모델링을 사용하여 분야별 연구 동향을 파악하고자 하는 연구가 진행되고 있다[21]. 회계학에서 토픽모델링 기법을 사용한 연구동향 연구는 아직까지 찾아보기 힘들다.

논문의 초록을 분석하여 주제들을 추출하고 저자를 식별하거나 시간의 흐름에 따른 주제 분포의 변화를 통해 해당 학문분야의 연구 동향을 파악하는 연구들에서 토픽모델링 기법이 활발하게 적용된다[19]. 이러한 연구들은 공통적으로 학술지나 박사학위 논문을 대상으로 하여 논문의 초록에서 10개 이내의 토픽을 추출하고 그 결과를 분석한다. 데이터의 수집과 분석에는 R이나 파이썬(Python)과 같은 프로그램이 사용되며, 결과의 해석에는 키워드 빈도 분석, 토픽 트렌드 분석 등의 방법이 이용된다.

문서 집합을 구성하는 단어들을 분석하는 키워드 빈도 분석은 토픽모델링 기법을 사용하는 연구에서 공통적으로 수행되는 부분이다. 수집한 초록 데이터에서 단어들을 추출하여 각 단어의 출현 빈도를 파악한다. 이 단계에서 문서 간의 차이를 구별해 주는 단어의 가중치를 반영한 TF-IDF (Term Frequency - Inverse Document Frequency)를 이용하여 단어의 중요도를 평가하기도 하고[22], 워드클라우드(Word Cloud)를 이용하여 단어의 출현 빈도를 시각화하기도 한다[21,23]. 워드클라우드란 단어의 출현 빈도에 가중치를 두어 글자의 크기 등을 다르게 나타냄으로써 텍스트 데이터를 시각적으로 표현하는 방법으로, 특정 주제를 직관적으로 설명하는데 유용하다. 연구자의 판단에 따라 전체 데이터를 대상으로 하거나[23,24], 데이터를 여러 그룹으로 나누어 실행함으로써[21] 시각적인 비교 가능성을 높일 수 있다.

시간의 흐름에 따라 토픽이 변동하는 경향을 파악하는 토픽 트렌드 분석은 시대별로 어떤 토픽이 주된 연구주제인지를 확인할 수 있어 토픽모델링을 활용한 연구에서 결과의 분석 방법으로 사용된다. 황은희 등(2019)은 자유학기제 연구를 정책의 도입, 실행, 안정화 및 확산 단계

로 나누어 각 시기에 따른 토픽의 등장 확률을 분석하고 시기별 연구의 변화 동향을 파악하였으며, 이를 바탕으로 정책이 추진되고 시행되는 추진 과정 및 시기별 쟁점에 따라 관련 연구의 주요 관심사도 함께 변화하고 있음을 확인하였다[22]. 조경원 등(2019)은 4차 산업 관련 연구 동향을 분석하는 과정에서 토픽의 변화 추이를 파악하여, 초창기에 활발했던 인문학적인 논의는 연구 측면에서 감소 추세이고 사물인터넷(IoT)에 대한 논의는 2019년에 크게 증가하였음을 확인하였다[23].

그 외에 빈도 분석에서 파악된 단어들로 상호간의 연관 관계를 파악하는 키워드 연관성 분석이나[11,25], 토픽의 비중과 토픽 간 거리를 나타내는 그림을 통해 토픽 간의 연관성과 유사도를 파악하는 방법도 결과의 해석 방법으로 사용된다[23].

3. 연구방법

3.1 데이터의 수집과 전처리

본 연구의 분석대상 논문은 KCI[26]에서 대분류 사회과학, 중분류 회계학 또는 경영학으로 분류되는 학술지에 2001년부터 2019년까지 게재된 논문 중 “이익”과 “조정”의 키워드를 이용해 AND 조건으로 검색되면서 한국어 초록이 존재하는 1,454건의 논문이다. 파이썬의 웹스크래핑 기법을 이용하여 분석대상 논문의 고유 URL에 대응되는 서지정보를 순차적으로 방문하여 한국어 초록을 읽어오도록 프로그래밍 하였으며, 그 결과 1,454건의 데이터가 수집되었다. 분석도구로는 파이썬 3.6을 이용하였다.

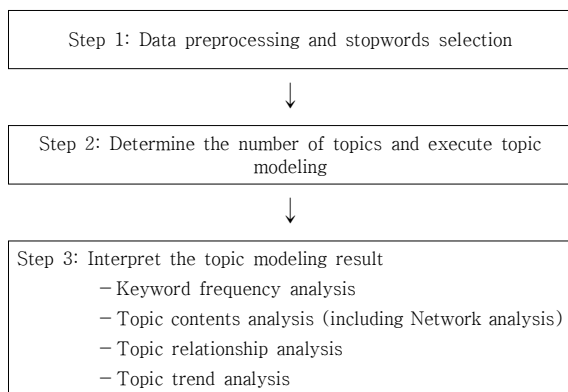


Fig. 1. Research Process

본 연구는 파이썬을 이용하여 수집된 초록 데이터에 대해, 데이터의 전처리와 불용어의 설정, 토픽 수의 결정과 토픽모델링의 실행, 결과 해석까지 총 3단계에 걸쳐 연구를 진행하였다. 본 연구의 전체 흐름은 Fig. 1과 같다.

전처리란 데이터를 분석하는 작업 전에 데이터를 분석하기 좋은 형태로 만드는 과정이다. 데이터 분석 과정에서 데이터 전처리가 중요한 근본적인 이유는 아무리 좋은 분석 기법이나 도구를 사용하더라도 품질이 낮은 데이터로는 좋은 분석 결과를 얻을 수 없기 때문이다[27]. 본 연구에서는 수집된 데이터를 정제한 뒤 전처리하여 명사만을 추출하였고, 그 결과 4,901개 단어가 중복을 포함하여 총 262,400회 등장하는 것을 확인하였다. 그 뒤 최빈도 단어 200개에 대해 IDF분석을 실행하여 IDF 값이 1 미만인 28개 단어를 불용어로 설정하였다. IDF(Inverse Document Frequency)는 특정 단어가 전체 문서 중 몇 개의 문서에 등장하는지(Document Frequency)를 측정하는 개념이다. 상대적으로 많은 문서에 출현한 단어의 IDF 값은 작아지고, 반대로 특정 문서에 편중하여 나타난 단어의 IDF 값은 커진다[28]. 최빈도 200개에 속하지만 IDF 값이 낮은 단어는 문서집합 전체에 자주 출현하지만 동시에 많은 개별 문서에도 등장하는 단어로, 문서집합 전체에서 보편적으로 사용되는 단어라고 해석할 수 있다. 본 연구에서 파악한 이러한 단어에는 문서집합의 주제가 이익조정에 관한 것임을 알 수 있게 해주는 ‘이익’, ‘조정’의 단어가 포함되었으며, 데이터의 원천이 논문임을 추측할 수 있게 해 주는 ‘연구’, ‘결과’, ‘분석’, ‘실증’ 등의 단어도 포함되었다. 논문을 작성할 때 논리적인 전개를 위해 많이 사용하는 ‘첫째’, ‘둘째’, ‘또한’ 등의 단어도 여러 문서 걸쳐 자주 등장하는 단어로 확인되었다. 불용어 설정 후 단어 집합에는 4,873개 단어가 남았으며 이 단어들의 출현 횟수는 중복을 포함하여 총 185,058회이다.

3.2 데이터의 분석과 결과해석

연구의 두 번째 단계는 토픽 수의 결정과 토픽모델링의 실행 단계이다. 토픽모델링을 위해서는 모델을 실행하기에 앞서 토픽 개수를 결정하여야 한다. 다만, 적절한 토픽 수가 무엇인지에 대한 통계적인 해법은 없으며 도출된 토픽들의 해석가능성과 타당도 및 연구 질문에 따라 좌우되므로, 토픽 수는 연구자들이 토픽모델링을 통해 유의미한 결과를 도출하는데 가장 적절하다고 판단한 숫자로 결정된다[19]. 선행 연구에서는 연구자의 판단에 따라

10개 이내의 토픽 수를 결정하여 토픽모델링을 실행하였으나, 본 연구에서는 토픽 수 결정에 연구자의 주관이 개입되는 것을 최소화하고자 토픽 일관성 값(Topic Coherence Value)을 기반으로 토픽 수를 도출하였다. 토픽 개수가 2개에서 20개까지 하나씩 변화할 때마다 각각의 토픽 수에서 토픽 일관성 값을 측정하는 함수를 파이썬에서 구현하였으며, 그 결과 Fig. 2와 같이 토픽 수가 4개일 때의 모델에서 토픽 일관성 값이 0.2966로 가장 높게 나타나는 것으로 확인되었다.

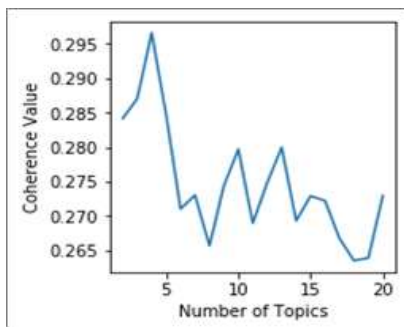


Fig. 2. Coherence value as a function of number of topics

토픽모델링을 실행하기 위해서는 코퍼스(corpus), 사전(id2word) 그리고 토픽 개수(num_topics)를 지정하여야 한다. 본 연구에서 코퍼스는 185,058회의 단어 출현이고 사전은 4,873개의 유니크한 단어이다. 모델에 적용된 토픽 수는 4개이다. 본 연구에서는 이렇게 설정한 정보를 바탕으로 Fig. 3과 같이 반복 횟수가 1,000회인 LDA 토픽모델링을 실행하였다.

```
model = models.ldamodel.LdaModel(corpus, id2word=id2word,
num_topics=4, random_state=100, iterations=1000)
```

Fig. 3. LDA topic modeling code in Python

본 연구에서는 선행 연구를 참고하여 키워드 빈도 분석, 토픽의 내용 분석, 토픽 간의 관계 분석, 토픽 트렌드 분석의 방법을 이용하여 토픽모델링 결과를 해석하였다. 우선 키워드 빈도 분석에서는 토픽모델링의 대상이 되는 4,873개의 단어에 대해 출현 빈도가 높은 키워드 20개를 추출하고 각각의 빈도수를 확인하였다. 그리고 단어 전체에 대해 워드클라우드를 실행하여 단어의 출현 빈도를 시각적으로 확인하였다. 이후 4개의 토픽에 대해 각 토픽별로 주요 키워드를 추출하고 토픽별로 배분된 논문 내용의 내용을 검토하여 각 토픽이 어떤 내용으로 구성되는

지를 확인하는 토픽 내용 분석을 수행하였다. 이 과정에서 각 토픽에 대해 지지도를 0.08로 하는 네트워크 분석을 추가로 수행하여 단어들 간의 연관성 분석에 참고하였다. 토픽 간의 관계는 파이썬 pyLDAvis 라이브러리를 이용한 시각화 자료를 토대로 분석하였으며, 마지막으로 시간의 흐름에 따라 토픽이 변화하는 내용을 확인하여 토픽 트렌드의 특이사항에 대해 그 원인을 분석하였다.

4. 연구결과

4.1 키워드 빈도 분석

2001년부터 2019년까지의 이익조정 관련 1,454편의 연구에서, 데이터를 전처리하고 불용어를 설정한 후 최종적으로 선정된 4,873개의 단어는 중복을 포함하여 총 185,058회 출현한다. 등장 빈도가 높은 순서대로 1위부터 20위까지를 확인한 결과는 Table 1과 같다.

Table 1. Top 20 keywords by frequency

Rank	Word	Freq.	Rank	Word	Freq.
1	Auditing	4,142	11	Introduction	1,020
2	Finance	1,195	12	Measure	1,008
3	At one's will	1,192	13	Value	1,006
4	Reduction	1,174	14	Market	1,001
5	Operation	1,152	15	Prediction	954
6	Effect	1,139	16	Asset	950
7	Relevance	1,119	17	Variable	948
8	Cost	1,113	18	Existence	945
9	Level	1,077	19	Possibility	932
10	Evaluation	1,076	20	Definition	907

Table 1에서 확인할 수 있듯이 ‘감사’가 4,142회로 가장 많이 등장하였으며, ‘재무’가 1,195번, ‘자의’가 1,192번, ‘감소’가 1,174번으로 그 뒤를 이었다. 키워드 빈도 분석을 통해 이익조정 관련 연구가 전반적으로 감사와 연관되어 있음을 짐작할 수 있으며, 회사가 자의적으로 재무나 영업과 관련된 비용을 감소시키거나 어떤 것의 도입을 통해 시장에서의 가치를 변화시키고 무엇인가를 예측하고자 하는 내용이 이익조정 연구에서 논의되고 있다는 것을 파악할 수 있다.

전체 단어에 대해 워드클라우드를 실행한 결과는 Fig. 4와 같다. 이를 통해 출현 빈도가 가장 높은 ‘감사’가 가장 큰 비중을 차지하며 ‘재무’, ‘자의’ 등의 단어도 자주 등장하는 단어임을 시각적으로 확인할 수 있다.



Fig. 4. Result of word cloud

4.2 토픽의 내용 분석

Table 2는 토픽모델링 실행 결과 4개의 토픽에서 각각 주요 키워드로 도출된 10개의 단어를 출현 빈도순으로 나열한 것이다.

Table 2. Top 10 keywords by topic

Rank	Topic 1	Topic 2	Topic 3	Topic 4
1	Auditing	Corporate Tax	Auditing	Variable
2	Operation	Transaction	Finance	Asset
3	At one's will	Prediction	Introduction	Investment
4	Effect	Cost	Relevance	Debt
5	Quality	Debt	Reduction	Auditing
6	Fee	Ratio	Cost	Market
7	Existence	Market	Evaluation	Cost
8	Activity	Value	Prediction	Effect
9	Introduction	Auditing	Level	Evaluation
10	Level	Income	Model	Financial Statements

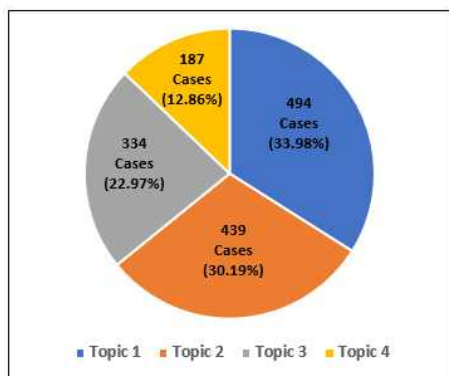


Fig. 5. Number of papers by topic

Fig. 5는 분석대상 논문이 각 토픽에 배분된 결과이다. 분석대상 논문 1,454건 중 494건(33.98%)이 토픽1

로 분류되었으며, 439건(30.19%)은 토픽2로, 334건(22.97%)은 토픽3으로, 187건(12.86%)은 토픽4로 분류되었다.

토픽1로 분류된 문서들은 감사 보수, 감사인 교체, 감사 시간, 감사보고서 감리, 감사보고서 발행, 감사 의견 표명 등과 관련된 이익조정 연구가 주를 이룬다. 주요 단어 30개 중에는 ‘감사’, ‘품질’, ‘교체’, ‘보수’가 토픽1에 편중되어 배분된 것으로 확인되는데, 이 중 ‘품질’은 ‘감사 품질’에서, ‘교체’는 ‘감사인 교체’에서 주로 등장하며, ‘보수’는 ‘감사 보수’를 명사별로 토큰화 하는 과정에서 분리된 것으로 나타난다. 또한 네트워크 분석 결과 토픽1의 단어들은 Fig. 6과 같이 ‘감사’를 중심으로 강하게 응집되는 모습을 보인다. 이상의 분석을 종합하여 볼 때 토픽1은 감사 및 감사보고서를 주된 내용으로 하는 토픽인 것으로 해석된다.

토픽2로 분류된 문서들은 타인자본과 자기자본 비율에 따른 이자비용 차이의 법인세 효과, 이연법인세 자산 및 부채와 관련된 이익조정, 법인세율 변동에 따른 이익조정 등을 연구 주제로 다루고 있다. 주요 단어 30개 중에는 ‘법인세’, ‘거래’, ‘소득’, ‘특수’, ‘관계자’ 등의 단어가 토픽2에 편중되어 배분된 것으로 나타난다. 이 중 ‘특수’와 ‘관계자’로 구성된 ‘특수관계자’는 세무 분야에서 일반적으로 많이 사용되는 용어이며 특히 법인세 분야의 부당행위계산부인과 밀접하게 연관되어 있는 단어이다. 이러한 이유로 ‘특수’와 ‘관계자’가 ‘법인세’를 주요 키워드로 하는 토픽2로 분류된 것으로 보인다. 토픽2에 포함된 문서 및 단어를 종합하여 판단해 볼 때, 토픽2는 기업의 부채비율이나 법인세 등이 기업가치에 미치는 영향 및 이를 이용한 경영전략을 주된 내용으로 하는 토픽인 것으로 해석된다.

토픽3으로 분류된 문서들은 원가구조, 기업지배구조, 기업의 국제다각화, 투자기간, 예측정보, 기업 전략 등과 관련된 이익조정 연구를 다루고 있다. 토픽3의 주요 단어 30개에는 다른 토픽의 키워드로 등장하지 않은 ‘모형’, ‘미래’, ‘매출’, ‘수익’ 등의 단어가 포함되었는데, 그 중 ‘미래’와 ‘매출’의 토픽3 편중도가 다른 단어에 비해 높은 편이다. 데이터를 추가적으로 검토한 결과 ‘미래’의 경우 ‘예측’, ‘추정’, ‘성과’, ‘수익’ 등의 단어와, ‘매출’은 ‘현금’, ‘예측’, ‘감소’ 등의 단어와 관련 있는 것으로 나타났다. 단어들을 조합해 보면, 미래의 성과 또는 미래의 수익을 예측하거나 추정하는 내용이 토픽3에 존재하는 것으로 보이며 매출 증감에 대한 부분도 토픽3에서 다루어지고 있는 것으로 해석된다. 이상의 분석을 종합할 때, 토픽3은

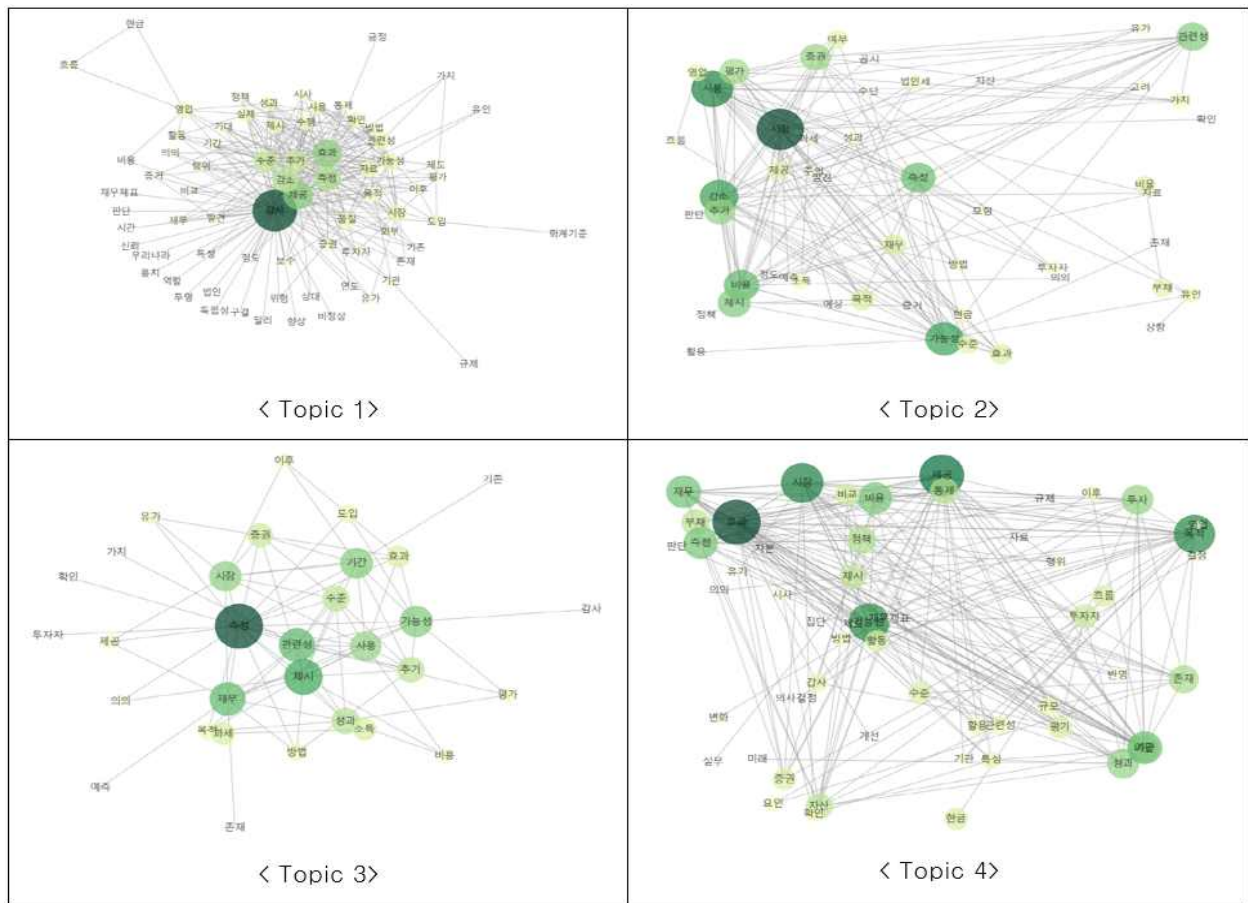


Fig. 6. Network analysis by using the word set belongs to each topic

기업의 전략적인 의사결정을 바탕으로 미래의 영향을 추정하고 그것의 적용 가능성을 평가하는 등, 전반적으로 기업의 경영전략과 관련된 것을 주된 내용으로 하는 토픽인 것으로 분석된다.

토픽4로 분류된 문서들은 재무제표의 표시, 보고양식의 선택, 회계정보의 조정, 감사인 지정 등과 관련된 이익 조정 연구를 다루고 있다. 토픽4에 편중되어 배분된 것으로 확인되는 단어는 '지정'이며, 데이터를 추가적으로 검토한 결과 '지정'의 연관 단어는 '감사', '부실', '관리종목', '잠재', '해제' 등이다. 또한 토픽4는 '재무제표'라는 키워드가 상위 10개 단어에 포함되었는데, 이는 재무제표를 토대로 잠재적으로 부실 가능성이 있는 기업을 판단하여 이를 관리종목으로 지정하거나 감사인을 지정하는 등의 내용과 연결되는 것으로 보인다. 토픽4에 포함된 문서 및 단어를 종합하여 판단해 볼 때, 토픽4는 재무제표와 회계 기준을 주로 연구하는 토픽이라고 할 수 있으며, 이를 통해 기업의 재무정보를 이용하는 이해관계자들에게 유용한 정보를 제공하는 방안을 연구하는 것을 주된 내용으로 하고 있는 것으로 나타난다.

이상의 분석을 토대로 토픽의 이름을 정리하면 Table 3과 같다.

Table 3. Topic name

	Topic name
Topic 1	Auditing and audit reports
Topic 2	Corporate taxes and debt ratios
Topic 3	General management strategy
Topic 4	Financial statements and accounting principles

4.3 토픽간의 관계 분석

토픽모델링 결과는 파이썬 pyLDAvis 라이브러리를 이용하여 시각화 할 수 있다. pyLDAvis에서 각 원은 토픽을 나타낸다. 원의 크기는 총 단어 등장 횟수 중 해당 토픽에 배분된 단어의 등장 횟수를 의미하며, 원이 클수록 그 토픽이 연구자들에 의해 많이 다루어지고 있는 주제라고 해석할 수 있다. 이익조정 연구의 토픽모델링 결과를 시각화 한 것은 Fig. 7과 같다.

Fig. 7에 의하면 토픽1의 크기가 가장 크고 그 다음이



Fig. 7. LDA topic modeling result visualization by topic

토픽2이다. 토픽1과 토픽2의 원 크기는 비슷하다. 이것은 이익조정을 연구하는 연구자들 사이에서의 주된 관심사가 토픽1과 토픽2임을 나타낸다. 토픽3과 토픽4는 상대적으로 원이 작다. 그러나 그 크기를 pyLDAvis 차트 하단의 한계 토픽 분포(Marginal Topic Distribution) 기준선과 비교해볼 때 두 토픽의 절대적인 비중은 각각 전체 연구의 10% 이상인 것으로 확인된다. 이를 해석하면 토픽3과 토픽4의 경우 토픽1과 토픽2에 비해 상대적으로 적게 다루어지기는 하나, 해당 토픽 역시 이익조정 세부 주제로 연구자들 사이에서 꾸준히 연구되고 있는 주제인 것으로 분석된다. 또한 4개의 원은 겹치는 부분 없이 서로 간에 약간의 거리를 두고 있는데, 이처럼 토픽들 간의 겹치는 영역이 없거나 작을 경우 각 토픽들이 비교적 명확한 연구 영역으로 구분되고 있음을 의미한다[23]. 이상의 분석을 종합하면 이익조정과 관련된 연구는 토픽1과 토픽2가 세부 연구 주제로 많이 다루어지고 있으나, 특정 토픽에 치우치지 않고 4개의 세부 주제가 고유의 영역에서 연구되고 있다고 할 수 있다.

4.4 토픽 트렌드 분석

4.4.1 논문 수를 기준으로 분석한 토픽 트렌드

2001년부터 2019년까지 이익조정 관련 논문의 연도별 토픽 트렌드를 논문 수(단위: 건) 기준으로 그래프로

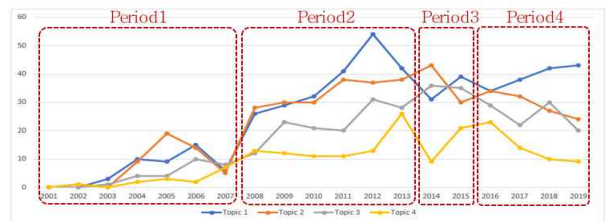


Fig. 8. Topic trend from 2001 to 2019 based on the number of papers

나타낸 것은 Fig. 8과 같다. 이익조정과 관련된 연구에서 분석대상 기간 전체에 걸쳐 연구자들의 가장 많은 관심을 받은 주제는 ‘감사 및 감사보고서’를 연구하는 토픽1이며, 대상 기간 동안 가장 적은 관심을 받은 주제는 ‘재무제표와 회계기준’을 연구하는 토픽4인 것으로 확인된다. 분석대상 기간을 논문 수의 흐름이 변동되는 기준으로 다음과 같이 4개의 기간으로 나누었다.

- (기간 1) 2001년부터 2007년까지: KCI가 정식으로 운영되기 전의 기간
- (기간 2) 2008년부터 2013년까지: 각 토픽의 연도별 흐름이 안정화되는 기간
- (기간 3) 2014년부터 2015년까지: 지배적인 연구 주제를 특정하기 어려운 기간
- (기간 4) 2016년부터 2019년까지: 토픽1이 다시 주된 연구 주제가 되는 기간

(기간 1)은 토픽 트렌드를 분석하기에 데이터가 충분하지 않은 것으로 해석하였다. 이 기간의 토픽 트렌드는 향후 별도의 연구를 통해 보완을 할 필요가 있을 것으로 보인다.

(기간 2)로 구분한 2008년부터 2013년까지는 각 토픽의 연도별 흐름이 안정되는 기간으로, 대부분의 토픽에서 대체적으로 논문 수가 증가하는 경향을 보인다. 이는 이익조정 관련된 연구가 점차 늘어나고 있는 과정에서 연구자들의 관심이 특정 토픽에 집중되는 것이 아니라 전 토픽에 걸쳐 고르게 증가하고 있음을 보여준다. 이 기간에 논문 수가 크게 변동한 토픽은 토픽1과 토픽4인데 토픽1은 2012년에, 토픽4는 2013년에 급격히 논문 수가 증가하였다.

본 연구에서는 2012년에 토픽1의 논문 수가 증가한 이유에 대해, 다음의 세 가지를 원인으로 분석하였다.

첫째, 내부회계관리제도 모범규준 등의 개정이다. 내부회계관리제도는 기업이 경영목적을 달성하기 위하여 설치·운영하는 내부통제제도의 일부분으로서, 기업회계기준에 따라 작성·공시되는 회계정보의 신뢰성을 확보하기 위하여 기업내부에 설치하는 회계통제시스템이다 [29]. 주식회사의 외부감사에 관한 법률(이하 '외감법')에 따라 상장기업과 자산 1,000억 원 이상인 비상장 기업은 내부회계관리제도에 대해 2005년부터 외부감사인의 검토를 의무적으로 받고 있다(외감법 제8조). 금융당국은 2012년 12월에 한국상장회사협의회내에 설치된 내부회계관리제도운영위원회를 통해 「내부회계관리제도 모범규준」 등을 개정하고 이를 2013년 1월 1일부터 시행하였는데 [30], 내부회계관리제도는 외부감사인과 밀접한 연관이 있기 때문에 이 제도의 개정에 대한 관심이 감사를 주요 키워드로 하는 토픽1의 논문 수 증가로 연결된 것으로 보인다.

둘째, 감사인의 품질관리제도에 대한 감리실시 도입 5년을 맞아 그 성과를 분석하고자 하는 수요가 있었다. 금융감독원은 2013년 1월 언론을 대상으로 한 정례브리핑에서 2007년부터 도입한 회계법인의 품질관리제도에 대한 감리실시 5년을 맞아 그간의 성과에 대하여 분석한 자료(제목: 회계법인 품질관리감리제도 도입 5년의 성과 및 과제)를 배포한 바 있다[31]. 2012년에는 품질관리감리 도입 5년을 맞아 해당 제도에 대한 실효성을 점검하고자 하는 시도들이 있었으며, 이에 따라 관련 연구가 증가하여 2012년의 토픽1 논문 수를 증가시키는 요인으로 작용한 것으로 보인다.

셋째, 한국채택국제회계기준(K-IFRS)의 전면 도입에

따라 감사보고서감리 방식에 대한 변화의 필요성이 대두되었다. 국제회계기준의 도입은 기존의 규정중심(rule-based) 접근방법에서 원칙중심(principles-based) 접근방법으로 회계처리와 공시에 근본적인 변화를 초래하므로 재무제표 정보에 대한 감사 및 감사와 관련된 감리제도 역시 이에 맞춘 개정과 개선이 필요하다 [32]. 이에 따라 금융감독원에서는 2012년에 K-IFRS제도 하에서의 감사보고서 감리제도 운영 및 개선 방안에 대한 연구용역을 발주한 바 있다[33,34]. 요약하자면 2011년 K-IFRS 전면도입으로 인해 감사보고서 감리 개선 방안에 대한 연구자들의 관심이 증가하였고, 이러한 관심이 2012년 감사 및 감사보고서를 주요 키워드로 하는 토픽1의 논문 수의 증가로 이어진 것으로 해석된다.

2013년에 토픽4의 논문 수가 증가한 이유에 대해서는 K-IFRS의 전면 도입 이후 변경된 회계기준을 적용하는 과정에서 발생한 이슈사항 및 재무보고 오류 등이 2013년경에 많이 논의되었기 때문인 것으로 해석하였다. 금융감독원은 K-IFRS 도입 2년이 지난 시점에서 그동안의 성과 등을 살펴보고 향후 개선이 필요한 부분과 추진과제 등 감독방향을 안내하고자, 2013년 7월 언론을 대상으로 한 정례브리핑에서 K-IFRS의 도입 현황, 2년간의 정착 노력, K-IFRS 도입에 대한 평가, 향후 과제 등에 대한 분석 내용을 담은 자료(제목: 한국채택국제회계기준(K-IFRS) 도입 2년의 평가와 향후 과제)를 배포하였다[35]. 또한 회계정보 작성자와 이용자 등을 대상으로 하여 IFRS 적용 관련 이슈사항 등을 설명하고, 재무제표 작성·이용시 애로사항을 청취하기 위해 「IFRS 적용 이슈사항 간담회」를 2013년 7월 16일 개최하고, 간담회에서 논의된 내용을 요약 정리하여 언론에 배포(제목: IFRS 적용 애로사항 청취를 위한 간담회 개최 및 이슈사항 등 안내)한 바 있다[36]. K-IFRS의 전면 도입은 2011년이나 이 주제가 이익조정 연구 대상으로 많이 다루어진 것은 도입 2년 뒤인 2013년인데, 이는 해당 주제 관련하여 연구에 사용될 만한 충분한 데이터가 확보되는데 일정 기간이 소요되었기 때문으로 추정된다. 또한 2013년은 K-IFRS 적용과정에서 발견된 회계처리 이슈들을 공유하는 시기이고, 이러한 이슈 해결에 참고할 수 있는 통일된 해석 사례는 아직 만들어지기 전이다. 즉, K-IFRS 이슈 사항에 대한 다양한 의견과 다양한 사례가 존재하는 시기였기 때문에 이들을 이용한 연구도 활발히 진행되었을 것으로 추측되며, 이러한 배경으로 인해 2013년에 토픽4의 논문 수가 증가한 것으로 보인다.

(기간 3)으로 구분한 2014년부터 2015년까지는 지배적

인 연구주제를 특정하기 어려운 기간으로, 토픽1, 토픽2, 토픽3은 이 기간 동안 논문 수에서 우열을 가리기 어렵다. 지배적인 연구주제를 특정하기 어려우며, 해당 기간 동안 회계학에서 연구자들의 관심을 집중시킬 만한 사건들이 두드러지게 나타나지 않았다는 것을 유추할 수 있었다.

(기간 4)로 구분한 2016년부터 2019년까지는 토픽1이 다시 주된 연구주제가 되는 것을 확인할 수 있었다. 그 원인으로서는 해당 기간에 잇달아 이슈가 된 분식회계 논란 때문으로 분석된다. 2016년 대우조선해양 사태[37]에서 시작하여 2017년 한국항공우주산업(KAI)[38]과 2018년 삼성바이오로직스[39]까지 이어진 일련의 분식회계 논란이 감사 및 감사보고서를 키워드로 하는 토픽1의 논문 수 증가로 이어진 것으로 보였다.

4.4.2 논문 비중을 기준으로 분석한 토픽 트렌드

2001년부터 2019년까지 이익조정 관련 논문의 연도별 토픽 트렌드를, 해당 연도의 전체 논문 대비 각 토픽의 논문 비중(단위: %)으로 그래프로 나타낸 것은 Fig. 9 과 같다.

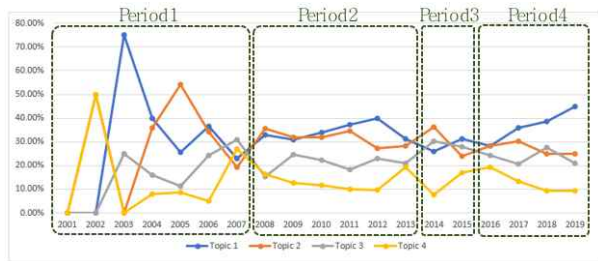


Fig. 9. Topic trend from 2001 to 2019 based on the percentage of papers

논문 비중을 기준으로 한 토픽 트렌드(Fig. 9)는 논문 수를 기준으로 한 토픽 트렌드(Fig. 8)와 유사한 흐름을 보인다. 기간1은 토픽의 연구 비중이 일관되지 못하고, 기간2는 토픽의 연도별 흐름이 안정화되면서 토픽1이 주된 연구주제로 나타난다. 기간3은 지배적인 연구 주제를 특정하기 어려우며, 기간4는 다시 토픽1이 주된 연구주제가 된다.

기간1을 제외하고 2008년부터 2019년까지(기간2부터 기간4까지)의 논문 비중 추이를 살펴볼 경우 4개의 토픽이 2개씩 매칭된다. 토픽1과 토픽4, 그리고 토픽2와 토픽3이 2014년을 제외하고 반비례 관계이다. 이는 Fig. 10 및 Fig. 11과 같이 토픽 비중 추이를 두 개씩 묶어서 비교할 경우 더 뚜렷하게 확인 가능하다.

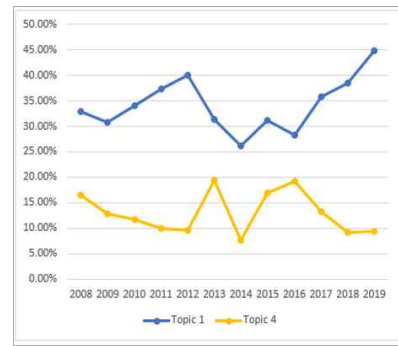


Fig. 10. Relationship between topic 1 and topic 4

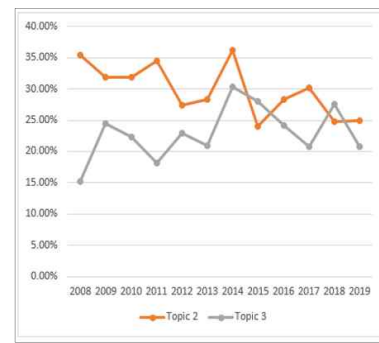


Fig. 11. Relationship between topic 2 and topic 3

본 연구자는 이러한 현상이 발생한 이유에 대해 이익조정을 연구하는 학자들이 크게 두 그룹으로 나누어지기 때문이라고 추측하였다. 첫 번째는 토픽1과 토픽4 관련된 이익조정을 주로 연구하는 그룹이고, 두 번째는 토픽2와 토픽3 관련된 이익조정을 주로 연구하는 그룹이다. 다만, 이 추정이 타당한지를 검증하기 위해서는 이익조정을 키워드로 하는 논문을 저자 기준으로 면밀히 분석하여야 할 것으로 보이며, 이 부분은 향후 별도의 주제로 연구될 수 있을 것으로 기대한다.

4.5 연구결과에 대한 토의

본 연구에서는 이전 회계학 분야 연구동향에 대한 권선국과 조성표(1997), 서은경과 한인구(1998), 권수영과 정종욱(1999), 최원석 등(2015)의 연구에서 보여준 바와 같이 이익조정 연구의 주된 토픽은 재무제표와 회계기준 일 것으로 추정하였으나, 실제 분석 결과 분석대상 기간 전반에 걸쳐 감사 및 감사보고서를 연구하는 토픽1이 주된 연구주제로 나타났다. 이를 토대로 다음과 같은 해석이 가능하다.

첫째, 감사가 이익조정의 통상적인 방지책으로 연구에

많이 언급되고 있다. 이익조정행위는 감사의 품질 향상과 범위 확대에 의해 제한될 수 있으므로 기업의 외부이해관계자는 기업 가치 평가 시에 감사품질을 중요한 요소로 고려한다[17]. 그러므로 이익조정과 관련된 연구에서 이익조정을 억제 또는 방지하려는 방안으로 감사가 많이 언급되었고, 그 결과 토픽1이 이익조정의 주된 주제로 도출되었을 수 있다. 둘째, 재무제표와 회계기준이 감사를 통해 간접적으로 이익조정에 영향을 미치고 있을 가능성이 있다. 회계기준의 신규 도입 또는 변경의 경우 이에 대한 재무적인 적절성을 검증하는 회계 감사도 변화에 맞추어 대응이 필요하기 때문이다. 2012년 토픽1의 논문 수 증가 요인으로 “K-IFRS의 전면 도입으로 인한 감사보고서 감리 방식에 대한 변화 논의”가 분석된 것에서 그 간접적인 영향을 유추해 볼 수 있다.

또한 논문 수 기준의 토픽 트렌드 분석에서 2011년부터 2013년까지 토픽1에서 토픽4까지 전반적으로 논문 수가 증가하였는데, 이는 2011년 K-IFRS가 전면 도입이 된 효과로 분석된다. 이익조정에서 재무제표와 회계기준 관련 토픽4의 논문이 가장 많이 생성된 시기는 K-IFRS가 전면 도입된 2011년일 것으로 예측하였으나, 해당 토픽에서 실제로 논문 수가 증가한 연도는 예상보다 2년 뒤인 2013년으로 분석되었다. 이는 새로운 회계기준 적용과정에서 발생한 회계처리 이슈들이 실제로 연구자들에 의해 발견되고 공유되는데 시간이 소요되었기 때문으로 추측되며, 이로 인해 연구에 사용될 만한 데이터가 확보된 시기가 2013년이었기 때문으로 해석된다.

5. 결론 및 시사점

본 연구에서는 텍스트 마이닝 기법 중 LDA 토픽모델링을 이용하여 ‘이익조정’을 연구 주제로 한 회계학 관련 논문들을 분석하여 해당 연구 주제와 관련된 세부 토픽에 대한 연구 동향을 탐색하였다. 2001년부터 2019년까지의 기간을 대상으로 하여 KCI에서 대분류 사회과학, 중분류 회계학 또는 경영학에 속하는 학술지에 게재된 1,454건의 논문을 분석 대상으로 하였다.

분석 결과, 이익조정의 세부 연구 주제는 감사 및 감사보고서를 주된 내용으로 하는 토픽1, 법인세 및 부채비율 등을 주된 내용으로 하는 토픽2, 기업의 전반적인 경영전략을 다루는 토픽3, 재무제표와 회계기준을 연구하는 토픽4로 나누어지는 것을 확인하였다. 또한 이익조정과 관련된 연구는 ‘감사’로 대표되는 토픽1과 ‘법인세’로 대표

되는 토픽2가 세부 연구 주제로 많이 다루어지고 있는 것을 시각적으로 확인하였으며, 4개의 세부 주제가 고유의 영역에서 연구되고 있다는 결론을 내렸다.

논문 수 기준의 토픽 트렌드 분석은 논문 수의 흐름이 변동되는 기준으로 4가지 기간으로 나누었으며, 각각의 기간에서 확인되는 트렌드의 특이사항에 대해 그 원인을 분석하였다. 2012년에 토픽1의 논문 수가 증가한 이유에 대해서는, 내부회계관리제도 모범규준 등의 개정, 감사인의 품질관리제도에 대한 감리실시 도입 5년을 맞아 그 성과를 분석하고자 하는 수요의 존재, K-IFRS의 전면 도입에 따라 감사보고서 감리 방식에 대한 변화의 필요성 대두를 원인으로 보았다. 2013년에 토픽4의 논문 수가 증가한 이유에 대해서는 2011년 K-IFRS의 전면 도입 이후 변경된 회계기준을 적용하는 과정에서 발생한 이슈 사항 및 재무보고 오류 등이 2013년경에 많이 논의되었기 때문인 것으로 해석하였다. 2016년부터 2019년까지는 해당 기간에 잇달아 이슈가 된 분식회계 논란 때문에 감사를 주요 키워드로 하는 토픽1의 논문 수가 증가한 것으로 분석되었다. 이어 논문 비중 기준의 토픽 트렌드 분석에서는 4개의 토픽이 2개씩(토픽1과 토픽4, 토픽2와 토픽3) 반비례 관계로 매칭되는 것을 발견하고, 저자 기준의 문헌분석을 향후 연구 과제로 제안하였다.

회계학분야 연구동향에 대한 과거 선행연구에서 권선국과 조성표(1997) 등 여러 연구들에 근거하면 이익조정 관련하여 재무제표와 회계기준이 주된 토픽일 것으로 판단되나, 실제 분석 결과에서는 감사가 이익조정과 가장 많이 연관되어 연구된 토픽으로 분석되었다. 또한 토픽 트렌드에서 2011년부터 2013년까지 이익조정 관련 연구가 전반적으로 높은 수준으로 증가하였는데, 이는 K-IFRS 도입이 영향을 미친 것으로 분석되었다. 재무제표와 회계기준 관련 토픽4에서는 K-IFRS이 도입된 2011년보다 2년 뒤인 2013년에 가장 많은 연구가 진행되었다.

본 연구는 다음과 같은 학술적 기여도를 갖는다. 첫째, 본 연구는 회계학 분야 연구동향에 대한 기존 문헌분석 연구에서 빈약하게 다루어 온 이익조정에 대해, 자세하게 연구 동향을 분석하여 정보를 제공하였다. 이익조정에 대한 시대에 따른 연구의 흐름을 파악하여 새로운 연구 분야를 발견하고 개척할 수 있게 하였다[8]. 둘째, 회계학의 중요한 연구 주제인 이익조정 관련 주된 토픽은 재무제표와 회계기준일 것으로 예측하거나 또는 경영자가 기회주의적 유인에 의하여 이익조정을 한다는 막연한 추측이 있었으나, 본 연구로 이익조정 관련 주요 토픽은 4가지로

구분되며 토픽1은 감사 및 감사보고서, 토픽2는 법인세 및 부채비율, 토픽3은 경영전략, 토픽4는 재무제표와 회계기준으로 구분된다는 것을 확인하였다. 셋째, 회계학 연구 동향을 분석하는 방법 중 하나로 텍스트 마이닝 기법을 제시하고, 기존의 문헌분석 연구의 한계점을 보완하였다는데 의의가 있다. 이를 통해 회계학과 경영학에 속하는 학술지 전체를 분석대상 학술지로 설정할 수 있었으며, 분석대상 기간을 KCI에서 제공하고 있는 자료의 최대 기간으로 넓게 설정할 수 있었다. 또한 데이터의 수집과 분석에 프로그램을 사용하여 자동화함으로써 인간의 실수가 발생할 가능성을 감소시키고 분석 과정에서 연구자의 주관적인 판단이 개입될 여지를 최소화시켰다.

본 연구는 실무적 측면에서도 다음과 같은 기여도를 갖는다. 첫째, 이익조정 관련 토픽트렌드를 프로그램을 통해 수치적으로 확인함으로써 실무적으로 정책 방향을 결정하는데 참고가 되는 정보를 제공한다. 불법적인 이익조정이 시장참여자에 의해 인지될 경우 기업은 투자자로부터 신뢰를 잃게 되어 기업가치는 심각하게 훼손될 수 있다[5]. 이런 이유로 금융당국은 회계기준 위반으로 인한 이익조정을 적발하고 예방하기 위해 지속적인 활동을 수행하고 있는데, 시장에 부정적인 영향을 미치는 이익조정을 사전에 방지하기 위해 정책결정자들이 회계기준 외에 감사 등 4개 토픽별로 추가적으로 고려하여야 할 사항과 그 추세에 대해 알 수 있도록 한다. 둘째, 기업은 이익조정 관련 4개의 토픽 유형을 참조하여 실무적인 이익조정 정책을 수립할 수 있을 것이다.

본 연구는 불용어 설정의 수준이 불명확하다는 연구의 한계가 있다. 앞서 분석 단계에서 최빈도 단어 200개 중 IDF 값이 1미만인 단어를 불용어로 설정하였는데, 이는 최적의 분석 결과를 도출해 내는데 부족한 수준이라는 의견이 제기될 수 있다. 한편 본 연구는 그룹1(토픽1과 토픽4)과 그룹2(토픽2와 토픽3) 내에서 이익조정 논문의 비중이 반비례 관계를 형성하는 것에 대해 연구자 기준의 분석을 향후 연구과제로 제시하였다. 이익조정을 키워드로 하는 논문들을 저자 기준으로 면밀히 분석할 경우 연구자들이 두 개의 큰 그룹으로 나뉘어져 있는지를 확인할 수 있을 것이며, 이를 통해 이익조정 연구에 대해 또 다른 의미 있는 정보를 제공할 수 있을 것이다.

REFERENCES

- [1] J. S. Han & M. S. Song. (2018). Trends and Future Role of Accounting Research. *Review of Accounting and Policy Studies*, 23(4), 1-32.
DOI : 10.21737/RAPS.2018.11.23.4.1
- [2] W. S. Choi, S. M. Yoon & S. R. Yoon. (2015). The Retrospect and Prospect on 40 Years of Accounting Research in Korea. *Korean Accounting Journal*, 24(1), 305-350.
- [3] S. K. Kwon & S. P. Cho. (1997). 20 Years of 『Korean Accounting Review』. *Korean Accounting Review*, 22(2), 203-220.
- [4] S. Y. Kwon & J. W. Chung (1999). Analysis of Research Trend in Korean Accounting. *Korean Accounting Journal*, 8(2), 19-53.
- [5] J. S. Choi, Y. M. Kwak & J. H. Back. (2012). The Association between Earnings Management and the Risk of CEO Turnover: A Survival Analysis. *Accounting & Auditing Research*, 54(1), 355-400.
- [6] C. B. Kim, C. S. Lee & H. S. Choi. (2007). A Theoretical consideration and Future Study on Earnings Management using Discretionary Accruals. *Korean International Accounting Review*, 18, 279-299.
- [7] T. K. Lim & S. W. Jeong. (2006). The Association between Discretionary Accruals and Analysts' Forecasts. *Accounting & Auditing Research*, 44, 311-334.
- [8] M. S. Kim, H. R. Jung, M. O. Kim & J. H. Park. (2015). Review of Management Accounting Case Study in Korea. *Korean Journal of Management Accounting Research*, 15(1), 71-112.
- [9] T. D. Kim & C. H. Bae. (2019). Review of Korean Case Research in Financial Accounting. *Korean Accounting Journal*, 28(2), 65-106.
DOI : 10.24056/KAJ.2019.03.002
- [10] E. G. Seo & I. G. Han (1998). Bibliographic Analysis of Accounting Studies. *Korean Accounting Review*, 23(3), 183-209.
- [11] S. A. Kim, J. H. Park, H. J. Lee & Y. J. Chung. (2016). A Study on the Research Trends of Multicultural Art Education Using the Text Mining Technique. *Multicultural Education Studies*, 9(2), 203-227.
DOI : 10.14328/MES.2016.6.30.203
- [12] T. M. Song & J. Y. Song. (2016). *Social Big Data Research Methodology with R*. Seoul : Hannarea Publishing Co.
- [13] J. W. Choi & Y. J. Lee. (2011). In-house Specialists in Accounting and Disclosures and the Accuracy and Reliability of Management Forecast. *Korean Accounting Journal*, 20(4), 157-184.
- [14] M. T. Kim, J. B. Wee & S. I. Jeon. (2006). Controlling Effects of Corporate Bonds Grading System on Earnings Management. *Asia-Pacific Journal of Financial Studies*, 35(5), 45-74.
- [15] I. M. Song, W. S. Paek & H. S. Park. (2004). Earning Management for the Avoidance of Losses. *Korean*

[1] J. S. Han & M. S. Song. (2018). Trends and Future Role

- Accounting Journal*, 13(2), 29–52.
- [16] H. D. Yi & K. J. Park. (2006). A Study on Earnings Management of Firms by Ownership Percentage Changes of Majority Shareholders. *Accounting & Auditing Research*, 43, 65–91.
- [17] J. K. Kim, S. M. Yu & G. O. Kim. (2011). The Effect of Audit Quality on Earnings Management. *Korean International Accounting Review*, 38, 43–74.
- [18] H. Y. Park, H. Y. Lee & M. J. Kang. (2012). The Impact of IFRS Adoption on Earnings management and Audit Hours. *Accounting & Auditing Research*, 54(2), 529–564.
- [19] C. H. Hahm. (2016). An Illustrative Application of Topic Modeling Method to a Farmers Diary. *Institute of Cross-Cultural Studies*, 22(1), 89–135.
- [20] D. Blei. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55(4), 77–84. DOI : 10.1145/2133806.2133826
- [21] H. I. Jo, J. W. Kim & B. G. Lee. (2019). A Study on Research Trends of Blockchain Using LDA Topic Modeling : Focusing on United States, China, and South Korea. *Journal of Digital Contents Society*, 20(7), 1,453–1,460. DOI : 10.9728/dcs.2019.20.7.1453
- [22] E. H. Hwang, J. H. Jang & H. G. Yang. (2019). An Analysis of the Research Trends in Free-Semester Using Text-Mining Techniques. *Journal of Education & Culture*, 25(3), 299–318. DOI : 10.24159/joec.2019.25.3.299
- [23] K. W. Choi & Y. W. Woo. (2019). Topic Modeling on Research Trends of Industry 4.0 Using Text Mining. *Journal of the Korea Institute of Information and Communication Engineering*, 23(7), 764–770. DOI : 10.6109/jkice.2019.23.7.764
- [24] S. M. Kim & Y. J. Kim. (2020). Research Trend Analysis on Living Lab Using Text Mining. *Journal of Digital Convergence*, 18(8), 37–48. DOI : 10.14400/JDC.2020.18.8.037
- [25] Y. G. Hyun, J. H. Han, U. Chae, G. H. Lee & J. Y. Lee. (2020). A Study On Technical Trend Analysis Related to Semantic Analysis of NLP Through Domestic/Foreign Patent Data. *Journal of Digital Convergence*, 18(1), 137–146. DOI : 10.14400/JDC.2020.18.1.137
- [26] Korea Citation Index. <https://www.kci.go.kr>
- [27] Y. H. Lee. (2018). *Data Preprocessing Using Python and SQL*. Seoul : Chaos Book.
- [28] S. J. Lee & H. J. Kim. (2009). Keyword Extraction from News Corpus using Modified TF-IDF. *The Journal of Society for e-Business Studies*, 14(4), 59–73.
- [29] *Financial Supervisory Service Financial Glossary*. Financial Supervisory Service. http://fine.fss.or.kr/main/fin_tip/dic/financedic.jsp
- [30] *Best Practice for Internal Accounting Management System(Revised in Dec. 2012)*. Financial Supervisory Service Accounting Portal. <http://acct.fss.or.kr/fss/acc/bbs/list.jsp?bbsid=1295497166723&url=/fss/ac/1295497166723>
- [31] Financial Supervisory Service. (Jan. 15th, 2013). *Achievements and Tasks over the 5 Years of the Accounting Firms' Audit Quality Control* Financial Supervisory Service Press Releases. http://www.fss.or.kr/fss/kr/promo/bodobbs_view.jsp?seqno=16474
- [32] S. Y. Kwon, Y. H. Lee & M. H. Hwang. (2011). A Way to Improve Audit Review by FSS as IFRS and New ISA are Introduce. *Korean Accounting Journal*, 20(3), 29–66.
- [33] C. W. Ra & M. J. Oh. (2012). A Study on Effective Audit Quality Management Operation Plan for Unlisted Corporations Subject to Consolidate Reporting. *FSS Research Report*. Financial Supervisory Service.
- [34] FSS Accounting Portal Research Report. (2012). Financial Supervisory Service. <http://acct.fss.or.kr/fss/acc/bbs/list.jsp?bbsid=1295497378262&url=/fss/ac/1295497378262>
- [35] Financial Supervisory Service. (Jul. 13th, 2013). *Achievements and Tasks over the 2 Years of applying K-IFRS*. Financial Supervisory Service Press Releases. [http://www.fss.or.kr/fss/kr/promo/bodobbs_view.jsp?seqno=16907&no=2&s_title=도입 2년&s_kind=title&page=1](http://www.fss.or.kr/fss/kr/promo/bodobbs_view.jsp?seqno=16907&no=2&s_title=도입%202년&s_kind=title&page=1)
- [36] Financial Supervisory Service. (Jul. 29th, 2013). *Meeting to Listen to Difficulties in Applying IFRS and Share the Issues*. Financial Supervisory Service Press Releases. http://www.fss.or.kr/fss/kr/promo/bodobbs_view.jsp?seqno=16991&no=60&s_title=ifrs&s_kind=title&page=3
- [37] H. Y. Kwak. (Jun. 22nd, 2016). *An Arrest Warrant Was Requested to the Former CFO of Daewoo Shipbuilding and Marine Engineering Due to the Trillions of Won of Accounting Fraud*. Kyunghyang Shinmun Corporation. http://news.khan.co.kr/kh_news/khan_art_view.html?artid=201606221825001&code=940301
- [38] J. H. Song. (Aug. 2nd, 2017). *KAI's Stock Price Plunged 16.5% Due to Police Accounting Fraud Investigation*. Seoul Economic Daily. <https://www.sedaily.com/NewsView/10JLGEV4N>
- [39] W. G. Kim. (Nov. 15th, 2018). *Securities & Futures Commission Has Concluded that Samsung Biologics Deliberately Committed Accounting Fraud*. Korea Business News CO., LTD. <http://www.wowtv.co.kr/NewsCenter/News/Read?articleId=A201811140601&t=NNv>

김 지 연(JiYeon Kim)

[정회원]



- 2006년 2월 : 서울대학교 경영학과 (경영학사)
- 2005년 제40회 공인회계사 시험 합격
- 2005년 10월 ~ 2011년 3월, 2013년 12월 ~ 2020년 6월 : 삼일회계법인 근무
- 2020년 7월 ~ 현재 : 삼일세무법인 근무

- 2020년 8월 : 고려사이버대학교 융합정보대학원(정보학 석사)
- 관심분야 : 빅데이터, 회계학, 경영전략
- E-Mail : dela8211@naver.com

나 흥 석(HongSeok Na)

[정회원]



- 1994년 2월 : 고려대학교 전산과학 학사
- 1996년 2월 : 고려대학교 이학석사(전산학 전공)
- 2004년 2월 : 고려대학교 이학박사(전산학 전공)
- 2002년 3월 ~ 현재 : 고려사이버대학교 소프트웨어공학과, 융합정보대학원 전임교수

- 관심분야 : 데이터베이스, 빅데이터, 인공지능, SW교육, 이터닝
- E-Mail : hsna99@cuk.edu

박 경 환(Kyung Hwan Park)

[정회원]



- 1997년 8월 : 충남대학교 경영학과(경영학박사)
- 2001년 3월 ~ 현재 : 고려사이버대학교 경영학부 전임교수
- 관심분야 : R&D관리, 빅데이터와 인공지능, 경영혁신
- E-Mail : parkh@cuk.edu