

소셜 네트워크 빅 데이터 관심이슈 추출을 통한 “홈택스” 이용현황 분석

황옥선*

〈요 약〉

2015년 “홈택스”는 새로운 개념의 통합 시스템으로, 사용자 니즈에 만족하기 위하여 수정·보완을 위한 다양한 분석이 필요하다. 따라서 본 연구에서는 소셜 네트워크를 활용하는 수많은 “홈택스” 사용자들의 빅 데이터를 활용하여 시스템에서 무엇에 관심을 가지고 있는지를 분석하였다. 빅 데이터 분석도구는 텍스트(Textom), 유씨넷(Ucinet), 넷드로우(NetDraw) 등의 프로그램을 활용하였다.

그 결과, 빈도분석에서는 ‘신청안내’, ‘연구’, ‘연말정산서비스안내’, ‘안내’ 등이 최상위 순위로 표시되었다. 연말정산 안내와 신청, 연구에 대한 관심도가 높았다. 전체네트워크 분석에서는 ‘신고’, ‘신청’, ‘금액’, ‘서비스’, ‘발급’, ‘사업자’, ‘조회’, ‘납부’, ‘접속’ 등이 중심에 있다.

이것은 “국세청 “홈택스”” 사이트에서 사업자가 어떻게 신고, 신청, 조회, 발급, 납부를 하는지에 대한 접속방법과 실제적인 세액 등에 집중적인 관심을 가지고 있다는 것이다. 중심성 분석과 CONCOR 분석에서는 홈페이지나 모바일에 의해서 “홈택스” 등록 및 신청, 신고 및 납부·조회 그리고 소득세 납부 금액과 증명서, 서류, 영수증, 포인트, 전화상담 등에 군집 간 연결성이 존재하였다. 국세청 입장에서는 이 단어들에 대한 조치가 필요하다.

〈주제어〉 “홈택스”, 빅 데이터, 중심성 분석, CONCOR 분석, 관심

* 한라대학교 글로벌비즈니스 학부 교수, yesek41@naver.com

I. 서 론

국세청 “홈택스(HTS)”(이하 “홈택스”라 함)은 조세의 종류에 따라 독립적으로 구분되어 있거나 부분적으로 규합된 형태로 구성되어 있었으나, 2015년부터는 통합된 “홈택스”로 전면 개편하여 활용되고 있다. 이 시스템은 현재 전 국민에게 일반화되어 있는 상태라고 할 수 있다. 시스템 앱의 월간 사용자는 2018년 기준 69만 명(채혜린, 2018)에 이르고 있다.

시스템 공급자는 제공한 내용에 대하여 사용자가 무엇에 관심을 갖는지와 어떠한 기능에 만족하는지 등을 고려하고 분석하는 것이 필요하다. 이에 따라 국세청은 “홈택스” 사용자의 요구와 만족에 부응하기 위하여 매년 사용자대상 설문지조사를 실행하였다. 그러나 국세청에 의한 설문지 구성은 사용자 스스로가 관심을 가지는 내용을 포함하지 않을 수 있다. 사용자는 통합 시스템에 대하여 국세청이 제공한 설문 내용 이외에 또 다른 내용에 대한 궁금증이 존재할 수 있다.

인터넷에는 수많은 사용자들의 다양한 관심사항이 나타나고 빅 데이터로 존재한다. “홈택스”는 인터넷에서 활용되고 있기 때문에 소셜 네트워크 사용자들의 의견을 분석하여 시스템에 반영하는 연구가 필요하다. 하지만 기존의 연구에서는 연구자가 구성한 설문지를 제한된 피험자 대상으로 통계 처리하여 실증적으로 검증하거나 실태조사 및 사례분석을 하였다. 이러한 분석들은 일반화의 한계가 존재한다. 최근에는 다양한 분야에서 수많은 피험자의 의견을 포함하는 빅 데이터에 의한 분석으로 실제적인 일반화에 활용되고 있다. Baggio and Cooper(2010)도 SNS 사용 지식과 정보내용이 유행성 확산모델로 전달 및 전이된다고 하였다.

통합된 “홈택스”는 사용자들에게 세금과 관련된 다양한 서비스를 제공하고 있다. 배춘호(2009)는 “홈택스”의 시행으로 국세통합시스템과 인터넷 웹사이트 연결에 의한 신고·고지·납부가 가능해졌고, 다양한 세무정보의 접근가능성이 향상되었다고 하였다. 현재 이 시스템은 많은 사용자들이 접속하여 활용하고 있기 때문에 빅 데이터로 존재하고 있다. 빅 데이터가 체계적으로 분석되어 사용자의 니즈에 부응하여 개선되어지면 만족감이 상승될 것이다. 그러나 “홈택스”에 관한 빅 데이터 연구는 매우 미흡한 실정이다.

Spencer et al.(2013)은 과학자들이 전통적으로 주요한 주제에 대하여 실증적 연구를 해온 방법이 비용과 공간적·시간적 범위의 한계가 존재하기 때문에, 전 세계의 대중적·경험적 정보 빅 데이터 자료가 신뢰할 수 있는 추정치를 제공한다고 하였다. 한경훈(2011)은 “홈택스”에서 사용자가 중요시 하는 콘텐츠 특성이 서비스 가치와 사용자 만족도에 영향을 미친다고 하였다. 사용자가 원하는 요구사항을 반영하여 개선하는 것이 필요하다는 것이다. “홈택스” 사용자들은 인터넷에서 수많은 필요사항을 나타낸다. 따라서 본 연구에서는 “홈택스”에 관한 소셜 네트워크로부터 사용자 관심 이슈 빅 데이터 내용을 도출하여 개선 사항을 분석하는 것이 목적이다.

본 연구의 구성은 다음과 같다. 제Ⅱ장에서는 국세청 “홈택스” 사용자 만족 변수에 관한 내용

과 주요 선행연구, 소셜 네트워크에 존재하는 “홈택스”관련 빅 데이터의 활용에 대하여 정리하고, 제Ⅲ장에서는 빅 데이터 자료 분석절차, 데이터수집 및 분석방법 등을 설명한다. 제Ⅳ장에서는 추출된 주요 단어에 대한 빈도분석, 전체 네트워크 분석, 중심성 분석, CONCOR 분석 등의 결과를 분석하고, 마지막으로 제Ⅴ장에서는 본 연구의 한계점, 향후 연구방향 등을 제시하면서 결론을 맺는다.

Ⅱ. “홈택스”(HTS) 서비스 시스템과 빅 데이터 적용

1. 국세청 “홈택스”(HTS) 사용자 만족 변수

“홈택스”는 국세청 내부의 시스템인 국세통합시스템(TIS ; Tax Integrated System)에 기반을 두어 2002년 11월 본격으로 납세자 및 세무대리인의 신고와 민원을 해결하기 위해 만든 시스템이다(정성모 외, 2017). 그러나 이 시스템은 잦은 설계변경에 따른 노후화로 시스템통합과 표준화에 어려움이 많고, 특히 세법 개정과 30 여종의 신규 사업발생시마다 시스템을 별도로 구축하게 되어 시스템 구조가 복잡해지고 정확한 세원관리의 한계와 유지보수 비용이 크게 문제가 되었다(정성모, 2015). 기획재정부는 변화된 세정환경에 부응한 효율적인 세원관리 납세 서비스 개선을 위해, 2011년 8월 12일 국세청의 국세통합시스템(TIS)의 전면개편 사업을 결정하고 총 사업비 2,302 억 원을 투입하여 2012년부터 2014년까지 차세대 국세통합시스템을 구축하고, 2015년부터 가동하고 있다(기획재정부, 2011).

과거의 국세통합 시스템이 사업별로 시스템이 만들어져 있다면, 현재의 국세통합시스템은 “홈택스”의 시스템에 성격으로 분류되어 진다. 예를 들어 조회/발, 민원증명, 신청/제출, 신고/납부, 상담/제보 등으로 구분되었다(정성모 외, 2017). 즉 과거에 사용하던 시스템은 세무정보 조회 신고 등이 중심이 되는 협의의 “홈택스”라고 할 수 있다. 이 시스템은 세법 개정으로 신규 사업 발생시 각각의 시스템을 별도로 구축하게 되어 시스템 구조가 복잡해지고 여러 시스템으로 분산되어 있는 국세행정시스템을 하나로 모아 관리해야 하는 상황 이었다(정성모, 2015). 개선된 현재의 “홈택스”는 최초부터 각각의 시스템이 하나로 통합되어 구축하여 운용되도록 설계된 시스템이다.

“홈택스”의 시행은 국세통합시스템과 인터넷의 연결에 의하여 웹사이트를 통한 신고·고지·납부가 가능해졌고 다양한 세무정보의 접근가능성이 향상되었다(배춘호, 2009). 그러나 이 시스템은 사용자 니즈에 대한 정기적이고 지속적인 조사를 시행하여 개선함으로써 사용자들이 편리하게 활용할 수 있도록 하는 것이 필요하다. “홈택스” 활용의 확대는 시스템에서 제공된 서비스 품질들이 사용자들에게 유의적인 의미를 갖게 하는가에 달려있다. 이에 따라 선행연구에서는 “홈택스” 품질에 관한 연구가 실행되었다. 주요 선행연구에서 확인된 서비스품질 요인은 다음과 같다.

〈표 1〉 “홈택스” 품질 요인에 관한 주요 선행연구

선행연구	요 인
정성모 · 변정희 · 김상현 (2017)	질의응답의 서비스, 정보의 충분성, 시스템의 편리성, 정보의 신속 · 보안성
이성욱 · 오상훈 · 김현수 (2015)	이용용이성, 유형성, 보안성, 정보성, 시스템신뢰성, 반응성
최동춘 · 박종원 · 차현수 (2014)	웹사이트 디자인요인, 웹사이트 정보제공요인, 웹사이트 커뮤니티요인, 웹사이트정보보안요인, 경험가치, 웹사이트 충성도
김명중 · 송민정 (2013)	정보 품질, 시스템 품질, 서비스 품질, 신뢰도, 사용편리성, 업무효율성
배춘호 · 서희열 · 배수진 (2011)	신뢰성, 편의성, 사용용이성, 경제성, 보안성
한경훈 (2011)	용이성, 유연성, 신뢰성, 주관적 규범, 자기효능감, 경제적 효익, 콘텐츠 가치성, 콘텐츠 최신성, 콘텐츠 업무관련성
임재희 · 신성식 · 양해면 (2007)	용이성, 유연성, 신뢰성, 정확성, 적시성, 유용성, 신뢰성, 응답성, 충분성
황옥선 (2005)	신속성, 신뢰성, 사용용이성, 경제성, 서비스, 심적 부담감, 보안성
심태섭 · 송인국 (2005)	시스템의 신뢰성, 제공정보의 충분성, 문의에 대한 서비스, 제공정보의 최신성, 시스템의 접근편리성, 시스템의 보안성, 정보출력의 편의성

〈표 1〉에서 보는 바와 같이, 선행연구에서는 “홈택스” 품질 요인들이 다양하게 검증되어 존재한다. 하지만, 2015년 이후에 구축된 통합된 “홈택스” 사용자 만족도 관련 연구는 정성모 외(2017)의 연구 외에 거의 존재하지 않고 있다. 통합된 “홈택스” 품질에 관한 사용자 만족도 연구가 필요하다.

정성모 외(2017)의 연구에서는 심태섭 과 송인국(2005) 연구와 동일한 내용변수를 사용하여 비교하였다. 그 결과, 통합된 “홈택스”가 편리성(로그인의 편리함, 사이트 접속의 편리함, 화면 전환 시 전환 속도, 최초 로그인 시간, 사용방법의 편리함, 이용 중 장애 발생정도, “홈택스” 이용가능시간의 편리성, 정보 출력 시 대기시간, 최초 회원가입의 편리함)과 질의응답 서비스(문의 사항, 문의의 편리성, 문의사항에 대한 답변의 신속성)에서 기존 “홈택스”보다 상대적으로 미흡한 것으로 나타났다. 이러한 이유는 “홈택스” 시스템이 통합된 초기이기 때문에 시행착오, 오류와 시스템 불완전성이 존재하고 있다는 것을 나타낸다. 그리고 유의적인 사용자 만족도 요소는 시스템의 편리성과 정보의 신속 · 보안성(정확한 정보의 제공정도, 명확한 의미로 정보를 제공하는 정도, 최신정보로 업데이트 되는 정도, 제공된 정보를 이해하기 쉬운 정도, 정보를 짧은 시간에 찾을 수 있는 정도, 이용목적에 맞는 정보의 제공정도, 신고내용에 대한 보안(유출방지) 정도, 서면인쇄 등 정보출력의 편리성, 필요한 시기에 정보를 제공하는 정도, 개인정보에 대한 보안(유

출방지) 정도) 등으로 나타났다. “홈택스” 이용 시 가장 중요한 것은 필요한 시기에 이용 목적에 맞는 최신정보로 업데이트된 정확한 정보를 짧은 시간에 명확한 의미로 이해하기 쉬운 편리한 정보가 제공되어야 한다. “홈택스”가 통합되면서 이러한 부분이 더욱 강화되어 사용자만족도에 유의한 영향을 미친 것으로 판단된다(정성모 외, 2017). 비유적인 요소는 질의응답의 서비스, 정보의 충분성 등으로 나타났다. 즉 문의사항에 대한 답변의 신속성 및 편리성, 화면구성, 다양한 정보제공 등은 미흡한 것이다.

이 연구는 이전의 시스템에 대한 연구보다는 실제적인 가치가 존재한다. 하지만, 연구대상이 세무대리인으로 한정되어 있으며, 설문대상이 서울과 경기지역의 212명이다. 이와 같이 선행연구에서는 “홈택스” 품질의 각 내용변수들을 연구자가 선정하고 통계 처리하여 몇 개의 요인으로 묶어 명명하였다. 그러나 각 연구는 요인명이 동일할지라도 구체적인 내용에 대하여는 서로 차이가 발생한다. 내용변수에 대한 각각의 효과가 다른 것이다. 사용자가 스스로 요구하는 내용을 분석한다면, 실제로 사용하는 사용자측면에서의 요구사항을 파악함으로써 더욱 실제적인 유용성을 가질 수 있을 것이다. 이에 따라 국세청은 통합된 “홈택스” 사용자에게 서비스품질에 관한 자체적인 실태조사를 실행하고 있다. 하지만 사용자 스스로가 관심을 가지는 내용에 대한 연구가 필요하다.

“홈택스” 사용자들은 매우 다양하고 광범위하다. 시스템 사용자들은 인터넷에 다양한 의견을 나타내고 있으며, 이 내용은 빅 데이터로 존재한다. 그리하여 시스템 사용자 대상이 확대되어 빅 데이터 자료에 대한 관심이슈가 무엇인지를 분석하는 연구가 더욱 유용할 것이다. 즉 우리나라 전체의 많은 “홈택스” 사용자들에 의한 빅 데이터에 의한 분석이 더욱 일반화에 적합할 것이다.

2. 소셜 네트워크 빅 데이터의 “홈택스”(HTS) 서비스 시스템 적용

현대는 인터넷에 의한 SNS, 블로그, 인터넷 뉴스, 동영상 등의 활용이 일반화되어 있는 시대이다. 인터넷 정보는 많은 사람들이 활용하고 있는 방대한 양의 데이터로 저장되어 “빅 데이터”라는 용어로 생성되고 있다. 빅 데이터는 대용량의 다양한 분야 데이터를 구성하고 있기 때문에 빅 데이터 자체를 분석하거나 다양한 데이터를 결합하면 유용한 패턴의 가치창출을 나타낼 수 있는 것이다(황옥선 · 민정기, 2017). Douglas(2012)는 빅 데이터가 실시간 생성하는 대규모의 데이터 텍스트, 사진, 동영상 등의 비정형 데이터로서 다양성의 특성을 포함하는 정보자산이라고 하였다. 따라서 포털사이트나 SNS, 블로그 등에서 축적되고 있는 수많은 내용들은 체계적 분석을 거치면 매우 유용한 정보로 활용될 수 있는 것이다. 그러므로 빅 데이터를 활용한 연구는 사람들의 실생활을 시시각각으로 접근한 내용을 포함하고 있기 때문에 실질적 결과를 나타낼 수 있다(황옥선, 2017).

빅 데이터는 디지털 환경에서의 대용량 데이터이다. 이 데이터는 일정기간 동안 네티즌들이 올린 글들에 대한 전수조사의 내용이 포함되기 때문에 체계적인 분석에 의한 일반적인 내용을

도출할 수 있다. 따라서 빅 데이터 분석은 기존에 시행되었던 인터뷰, 설문조사, 표본 추출 등에서 발생하는 문제점과 표본의 오차를 줄일 수 있고, 표본 오차에 따르는 결과의 왜곡도 줄일 수 있다(정명재 외, 2014). 즉 공급자는 사용자가 “홈택스”에서 필요로 하는 관심사항을 추정하여 시스템에 적용하는 것보다 사용자가 실제로 요구하는 내용을 갖추게 될 때 만족도가 높아진다.

빅 데이터는 사람 중심의 다양한 문제를 분석하는 연구에서 매우 유용하게 적용되고 있다(황옥선·민정기, 2017). 소셜 네트워크의 사용자들은 “홈택스”에 관한 많은 관심사항을 올려놓는다. 이 자료들이 빅 데이터 이다. 빅 데이터 분석은 사용자가 “홈택스”에서 어떠한 내용에 관심을 가지는지를 파악할 수 있고, 국세청은 공급자로서 관심이슈에 대한 시스템 개선에 노력하면 될 것이다. 하지만 기존의 연구에서는 빅 데이터 개념이 생성되지 않았고, 적용할 수 있는 분석방법이 미흡하였다.

소셜 네트워크의 빅 데이터는 체계적인 분석프로그램 도구에 의하여 다양한 분야에서 연구되고 있다. 장아름 외(2015)는 트위터 데이터를 노드엑셀 빅 데이터 프로그램에 적용하여 부산국제영화제의 글로벌 커뮤니케이션 네트워크 분석을 한 결과, 아시아권이 특정한 한류 스타의 영향에 의한 구전효과, 비아시아권이 K-POP 트윗 활동이 높다는 것을 도출하였다. 이수진과 전유나(2016)는 트위터와 넷마이너 프로그램을 활용하여 경기도 지역의 관광 주요 주제가 페스티벌과 캠핑, 레크레이션 등을 추출함으로써 이 분야의 콘텐츠 개발 및 활성화가 많은 사람의 관심을 불러일으킬 것으로 판단하였다. 아마르자르갈 다그와더르지 외(2020)는 베이커리 소비자 인식을 텍스트롬(Textom) 분석도구를 활용하였다. 그 결과, 소비자들은 브랜드, 맛 집, 위치, 리뷰, 카페, 재료, 커피, 건강, 가격, 디저트, 여행 등을 인식하고 있다고 하였다. 황옥선(2020)은 텍스트롬(Textom), 유씨넷(Ucinet), 넷드로우(NetDraw) 등의 프로그램에 상업교육 소셜 네트워크의 빅 데이터를 적용하였다. 그 결과 관심이슈는 특성화 고등학교의 취업과 진학 이원화 프로그램, 기업실무 관련 자격증 교육, 대학교와의 협약된 진학 연계, 관광, 비즈니스, 금융에 대한 정책, 중학교의 상업교육 등으로 도출되어 상업교육정책의 개선방향을 제시하였다.

기존의 연구에서는 “홈택스”에 관하여 연구자가 설정한 설문지를 대상으로 하였다. 그리고 피험자 표본은 설문지에 응답한 사용자에 한정되었다. 본 연구에서는 소셜 네트워크 사용자들 스스로가 관심을 가지는 내용을 대상으로 수많은 실제적인 의견을 나타내는 빅 데이터를 활용하여 결과를 도출한다. 이러한 연구 결과는 기존 연구보다 일반화에 더욱 접근하게 될 것이다.

III. 연구 설계

1. 빅 데이터 자료 분석절차

빅 데이터를 처리하는 과정은 데이터의 규모, 데이터간의 이질성 등과는 관계없이 비슷한 과

정을 거친다(Firat and John, 2013). 즉 1단계는 여러 데이터소스 중 분석목적에 적합한 데이터를 선택하고, 2단계는 데이터를 정제한 후에 패턴과 관계를 도출하는 데이터마이닝을 거치며, 3단계는 데이터 조직 및 해석의 과정을 실행한다(황옥선, 2017).

빅 데이터는 분석도구에 의하여 실행된다. 본 연구에서는 텍스트툼(Textom), 유씨넷(Ucinet), 넷드로우(NetDraw) 등의 분석도구를 사용한다. “홈택스” 관련 텍스트는 포털, 카페, 뉴스, 블로그 등에서 수집하였다. 데이터 추출을 위한 검색 키워드는 해당 도메인의 일반적인 단어를 선정하거나 도메인 전문가가 데이터 분석 목적과 검색된 데이터의 관련성을 고려하여 선택한다(김현정 외, 2015). 본 연구에서는 “홈택스”를 직접적으로 활용하는 사용자, 세무신고를 주요 업무로 하는 세무대리인, 세무회계 내용 및 활용에 대한 전문가 등을 선정하여 검색 데이터에 대한 정제 및 확정에 참여하도록 요청하였다. “홈택스” 데이터 추출을 위한 검색 키워드 참여자는 서비스 시스템 사용자 7명, 세무대리인 2명, 세무회계 교수 1명 등 총 10명을 대상으로 1차 및 2차 정제에 걸쳐 인터뷰하였다. 그 결과 1차적으로 추출을 위한 키워드는 “홈택스” 관련 빅 데이터 연구가 초기이기 때문에 “홈택스”로 선정하는 것이 타당하다고 결정하였다. 2차 정제는 1차 정제에 의한 키워드 중에서 정제 단어 숫자를 선정하였다.

키워드의 빈도 추출을 위한 양적인 분석은 텍스트툼(Textom) 프로그램 도구를 활용하였다. 데이터마이닝에 의한 의미 있는 관계성 분석은 유씨넷(Ucinet), 넷드로우(NetDraw) 분석도구를 활용하였다. 소셜 네트워크 분석을 위한 전문 프로그램으로는 Ucinet, Pajek, NetMiner, Social metrics, NodeXL 등이 존재한다. 유씨넷(Ucinet)은 현재까지 발전된 대부분의 소셜 네트워크 분석 이론에 대한 기능을 포함하고 있기 때문에 학술 연구에 많이 활용되는 장점이 있다(서정원과 김형중, 2019).

데이터마이닝은 조직에 축적되어 있는 대용량의 데이터로부터 자동 또는 반자동적인 방법으로 데이터에 숨겨 있는 의미 있는 패턴, 규칙, 관계를 찾아내는 과정이다(노규성 외, 2019). 데이터마이닝 과정은 개인, 집단, 사회의 관계를 네트워크로 파악하는 사회연결망 개념에 의하여 사회적 관계로 도출된다. 사회연결망은 개인 또는 집단의 네트워크를 나타내는 각 노드(node)들 간의 상호의존적인 관계(tie)에 의해 만들어지는 사회적 관계 구조를 말한다(황옥선, 2017).

2. 데이터수집 및 분석방법

데이터 수집은 2019년 6월 1일부터 2020년 5월 31일로 설정하였다. 자료는 빅 데이터 분석 도구 텍스트툼(Textom)을 활용하여 양적 추출을 하였다. 즉 “홈택스” 분야 중심단어는 구글의 웹문서, 뉴스, 페이스북, 트위터, 네이버의 웹문서, 블로그, 뉴스, 카페, 지식인, 학술정보전체 그리고 다음에서는 웹문서, 블로그, 뉴스, 카페 등에서 “홈택스”를 입력하여 추출을 시도하였다. 텍스트툼(Textom)에서는 각 포털에서 “홈택스”를 검색하여 출현빈도 순위에 따라 포털사이트의 각 부

문마다 1,000개가 1차 추출되어 모두 7,806개의 단어들이 도출되었다. 그 단어들은 거, 하, 받, 되, 알, 써야, 하는 것, 믿, 품, 첫걸, 고르, 간, 남, 년, 정, 듣, 넣 등의 의미 없거나 유사하게 중복되는 단어들을 조정하여 상위 200개의 단어로 2차 정제되었다. “홈택스” 단어들은 빈도추출에 의한 양적분석이 되고, 유씨넷(Ucinet), 넷드로우(NetDraw) 분석도구에 적용되어 사회연결망에 의한 질적 분석으로 실행되었다.

IV. 분석 결과

1. 주요단어 빈도분석 결과

소셜 네트워크 사용자들이 “홈택스” 관련 이슈에 대하여 많이 언급하는 주요단어를 도출하기 위하여 2차 정제 데이터 200개를 인터뷰하였다. 그 결과 주요 단어 숫자는 30개로 확정하였다. 주요단어 빈도 순위는 다음과 같다.

<표 2> “홈택스” 데이터의 상위 주요단어 빈도분석 결과

순위	단어	빈도	%	순위	단어	빈도	%
1	신청안내	114	14.23%	16	해외금융계좌	18	2.25%
2	연구	65	8.11%	17	적극힌트	17	2.12%
3	연말정산 서비스안내	64	7.99%	18	경우	16	2.00%
4	안내	46	5.74%	19	귀속	16	2.00%
5	사업자등록	42	5.24%	20	사이버시스템	16	2.00%
6	영향	40	4.99%	21	납세자	15	1.87%
7	원천징수	38	4.74%	22	홍보대사	15	1.87%
8	관할	28	3.50%	23	방법	14	1.75%
9	세금	28	3.50%	24	세금상식	14	1.75%
10	개선방안	27	3.37%	25	근로소득	12	1.50%
11	쌍커풀	24	3.00%	26	공정위	12	1.50%
12	근로 장려금	23	2.87%	27	납부	11	1.37%
13	장려금	22	2.75%	28	소득세	11	1.37%
14	신고	20	2.50%	29	편리	7	0.87%
15	확정일	20	2.50%	30	지속가능	6	0.75%

상기 <표 1>에서 보는 바와 같이, “홈택스” 관련 검색어 빈도수는 ‘신청안내’, ‘연구’, ‘연말정

산서비스안내’, ‘안내’ 등이 최상위 순위로 표시되었다. “홈택스”는 2002년 4월 “홈택스”를 개시하고(국세청 보도참고자료, 2005), 2015년 통합된 “홈택스”이 개통되었다. 그러나 소셜 네트워크에서는 아직도 “홈택스” 안내에 대한 관심도가 가장 높다는 것을 알 수 있다. 또한 ‘연구’가 최상위 순위라는 것은 “홈택스”를 활용하는 최종사용자 입장에서의 다양한 연구가 필요하다는 것을 나타내는 것이다.

주요단어로는 ‘사업자등록’, ‘영향’, ‘원천징수’, ‘관할’, ‘세금’, ‘개선방안’, ‘쌍커풀’, ‘근로 장려금’, ‘장려금’, ‘신고’, ‘확정일’, ‘해외금융계좌’, ‘적극힌트’, ‘경우’, ‘귀속’, ‘사이버시스템’ 등이 도출되었다. 이 결과는 사용자들이 “홈택스” 사이버시스템에서 주요 단어들에 대한 다양하고 쉬운 사용자맞춤형 힌트를 제공하여 개선하는 방안을 필요로 하는 것이다. 즉 사용자들은 신고 확정 일에 대한 귀속, 관할 세무서, 쌍커풀을 했을 경우 세금 영향, 각 종 장려금, 원천징수, 사업자 등록 등에 대한 세세한 내용에 대하여 보다 알기 쉬운 안내에 많은 관심이 있다는 것이다.

상위빈도에는 ‘납세자’, ‘홍보대사’, ‘방법’, ‘세금상식’, ‘근로소득’, ‘공정위’, ‘납부’, ‘소득세’, ‘편리’, ‘지속가능’ 등도 포함되어 나타났다. 이 결과는 납세자에게 직접적인 근로소득, 소득세 납부방법 등에 대하여 보다 지속가능한 방법의 홍보와 세금상식 등의 편의가 필요하다는 것을 나타낸다. 또한 수요자 입장에서는 홍보대사 활용, 세금상식 등의 편리한 정보, 납부에 대한 공정한 처리 등에 대해서도 많은 관심을 가지고 있는 것으로 파악된다.

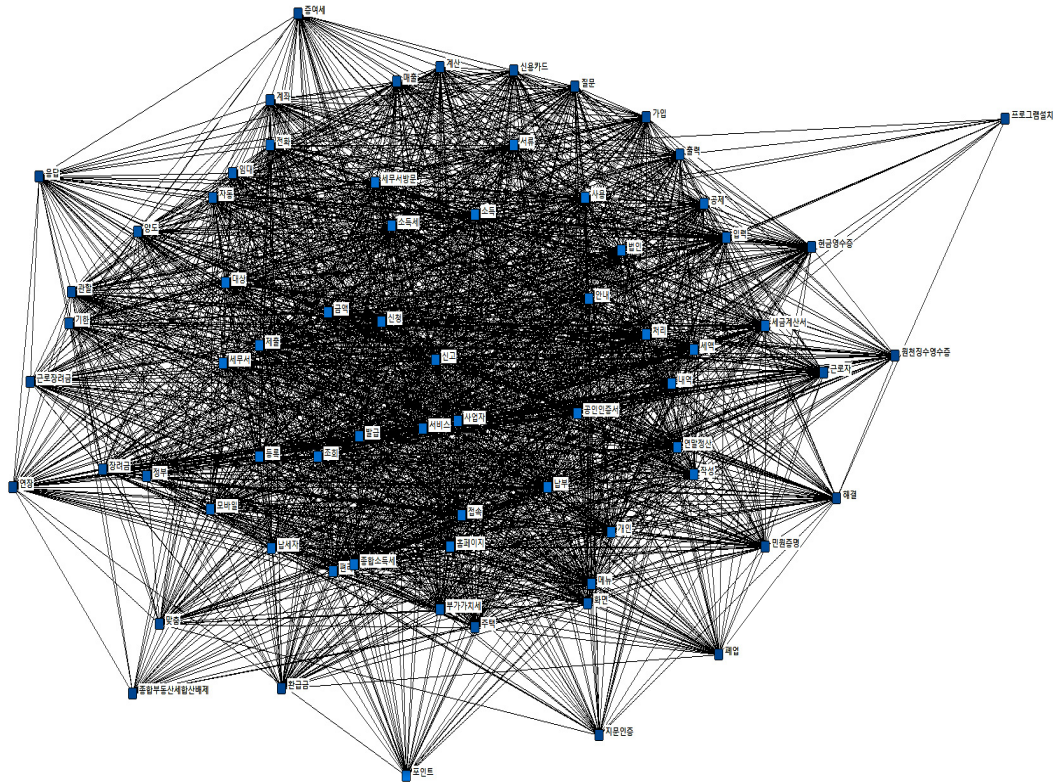
2. “홈택스” 관련 주요 키워드 전체 네트워크 분석

네트워크란 복수의 사람·조직·사물 등을 연결시키는 일정의 관계를 의미하고 있다. 네트워크를 형성하고 있는 사람·조직·사물 등을 액터(actor)라고 한다(21세기 정치학대사전). 본 연구에서는 “홈택스” 관련 주요 키워드 데이터들의 전체 네트워크를 시각화하여 액터(actor)들의 관계 분석을 하였다. 시각화 결과는 <그림 1>과 같다.

<그림 1>에서 보는바와 같이, 네트워크의 중심에는 ‘신고’, ‘신청’, ‘금액’, ‘서비스’, ‘발급’, ‘사업자’, ‘조회’, ‘납부’, ‘접속’ 등의 주요 단어들이 중심에 존재한다. 다양한 사업자들이 “홈택스”에 대한 신고, 신청, 조회 등의 서비스에 관심을 가지고 있다는 것을 나타낸다. 또한 금액은 얼마인지, 어떻게 납부하는지 등에 대한 관심을 가지고 있다. 그리고 사용자들은 가장 기본적인 접속에 대하여도 궁금해 하고 있는 것이다. 사업자들은 규모가 큰 사업자도 있겠지만, 소규모의 개인사업자들이 커다란 비중을 차지하고 있으며 “홈택스”에 관한 기초적인 지식을 가지고 있지 못할 수도 있다. 따라서 공급자인 국세청 입장에서는 알기 쉽게 찾아가는 교육서비스를 실행하여 “홈택스”에 대한 사용자 인터페이스가 원활하게 이용될 수 있도록 하는 것이 필요하다.

이와 같이 “홈택스”는 과거에 비하여 통합적이고 다양하게 개선되었지만, 빅 데이터 전체 네트워크의 중심부에 존재하는 이슈 사항을 검토해 볼 필요가 있는 것이다.

〈그림 1〉 “홈택스” 관련 주요 키워드 전체네트워크



3. “홈택스” 주요 단어 중심성 분석 결과

중심성은 사회적 네트워크분석에서 가장 대표적인 개념 중의 하나이다(Stephen, 2005). 중심성은 한 개체가 다른 개체들과 얼마나 많이 연결되어 있는지, 한 개체가 다른 개체들에 도달하려면 몇 단계가 필요한지 등을 측정할 수 있다(Cho, 2012). 중심성 척도 종류에는 어떠한 특정 노드에 연결된 노드들의 수를 양적으로 파악하는 연결정도 중심성(degree centrality)과 특정노드가 다른 모든 노드들에 가까운 정도를 파악해 주는 근접 중심성(closeness centrality), 특정 노드가 다른 노드들에 대하여 얼마나 중개역할을 하는지를 파악하게 하여 네트워크 내의 정보교환 또는 자원 흐름을 중개하는 경로상 위치하는 정도를 측정하는 매개 중심성(betweenness centrality) 등의 지표가 존재한다(서정원과 김형중, 2019 ; Choi 외, 2011). 그리고 본 연구에서는 연결정도중심성과 유사하게 노드들의 양적인 관계 나타내는 것이 아니라 연결의 질도 함께 반영하는 위세 중심성(eigenvector centrality)을 포함하였다. 이 척도는 중심적인 역할을 하는 행위자(액터)들과의 연결이 높으면 가중치를 주는 것이다(Wasserman and Faust, 1994).

주요 키워드에 대한 중심성 분석은 주제 내용을 고려하여 순위를 분석한다. 안명숙과 민용기(2016)는 1-14 순위와 1-21 순위의 2가지를 선택하였으며, 오익근 외(2015)는 1-20 순위를 선정하여 중심성의 비교 분석을 하였다. 본 연구에서는 1-20 순위까지의 키워드에 대한 분석을 하였다. 그 결과는 다음과 같다.

〈표 3〉 “홈택스” 데이터의 중심성 분석결과

주요 단어	연결정도	중심성 순위	근접 중심성 순위	매개 중심성 순위	위세 중심성 순위
신고	0.157	1	1	30.782	1
신청	0.091	2	0.986	3	18.533
사업자	0.076	3	0.945	11	13.175
발급	0.073	4	0.958	8	13.529
등록	0.071	5	0.986	3	18.533
연말정산	0.048	9	0.932	15	12.669
서비스	0.058	8	0.986	3	18.533
납부	0.064	6	1	1	30.782
조회	0.061	7	0.986	3	18.533
근로 장려금	0.029	18	0.775	20	3.677
모바일	0.044	11	0.958	8	13.477
소득	0.039	12	0.945	11	11.276
개인	0.045	10	0.972	7	17.533
공제	0.037	13	0.885	16	6.873
가입	0.031	14	0.885	16	8.722
편리	0.028	19	0.945	11	13.328
홈페이지	0.03	16	0.958	8	15.893
소득세	0.03	16	0.945	11	22.168
신용카드	0.031	14	0.831	18	3.362
세금계산서	0.017	20	0.831	18	3.779

상기 <표 3>에서 보는 바와 같이, 연결정도 중심성 순위에는 ‘신고’, ‘신청’, ‘사업자’, ‘발급’, ‘등록’ 등이 최상위를 나타내고 있다. 이 결과는 “홈택스” 실행 시에 사업자 등록을 하고 신고 및 신청 등의 탭을 클릭하여 시스템을 활용하는 것이 기본적인 절차이기 때문에 연결성이 높다고 할 수 있다. 또한 20 위권 내의 상위 순위에는 ‘연말정산 서비스’, ‘납부’, ‘조회’, ‘근로 장려

금’, ‘개인’, ‘공제’, ‘소득세’, ‘신용카드’, ‘세금계산서’ 등도 포함되어 있다. 이것은 연말정산의 근로 장려금, 소득세, 신용카드 및 세금계산서, 개인공제에 대한 조회와 납부 등이 다른 키워드들과의 연결성이 많다는 것을 나타낸다. 국세청은 편리한 연말정산 탭을 구성하고, 공제신고서 불러오기 등과 연결하여 예상세액 산출 탭을 구성하여 제시하고 있다. 그러나 사용자들은 그들의 입장에서 편리함을 요구하고 있기 때문에 소셜 네트워크에서 상위순위로 나타나고 있는 것이다. 따라서 이러한 단어들은 국세청 입장에서 보다 세심하고 알기 쉬운 구성에 우선적인 노력을 하는 것이 필요하다는 것을 나타낸다.

특히 사용자들은 신용카드에 의한 소득세 및 연말정산에 관심이 높다는 것을 알 수 있다. “홈택스”에서는 신용카드 소득공제 안내를 제시하고 있다. 그러나 세무대리인이 아닌 일반 사용자 입장에서는 그 내용에 대한 관심이 높으나 이해가 부담스러울 수 있다. 일반 사용자들은 신용카드 소득공제 보다는 직접적인 세액공제 정보가 필요할 것이며, 다양한 신용카드의 공제 공식 비율에 대한 이해가 부담스러울 수 있다. 따라서 국세청 입장에서는 사용자들이 신용카드에 의한 세금관련 정보를 쉽게 이해할 수 있도록 사용자 관점에서 신용카드 세액공제금액을 조회할 수 있도록 구성하면 더욱 유용할 것이다.

그리고 연결정도 중심성은 ‘홈페이지’, ‘모바일’, ‘가입’, ‘편리’ 등에서도 높게 나타났다. 근접 중심성에서는 ‘신고’, ‘납부’, ‘신청’, ‘등록’, ‘서비스’, ‘조회’, ‘개인’, ‘홈페이지’, ‘모바일’ 등의 단어가 서로 근접한 관계를 가지고 있는 것으로 나타났다. 이러한 결과는 사용자들이 홈페이지나 모바일에 의해서 “홈택스” 등록 및 신청, 신고 및 납부·조회 등에 관한 내용을 확인하는 경우가 매우 높다는 것을 의미한다. 또한 사용자들은 홈페이지나 앱의 가입에 대한 편리성에 관심을 가지고 있다는 것이다. 그러나 현재는 홈페이지와 모바일 앱의 화면이 다르게 구성되어 있기 때문에 사용자 입장에서 익숙하지 못하다. 사용자들은 두 화면에 접근해야 하는 번거로움이 존재한다. 따라서 국세청에서는 사무행정상의 원활한 사용자 인터페이스를 위하여 홈페이지와 “홈택스”에서의 신고/납부, 조회/발급 등에 대한 구성 탭의 동일성을 갖추는 것이 유용할 것이다. 이러한 결과는 정성모 외(2017) 등의 통합된 “홈택스” 시스템 초기 연구에서도 동일하게 나타났다. 사용자들이 지속적으로 관심을 가지는 사항이라고 할 수 있다.

매개 중심성에는 ‘신고’, ‘신청’, ‘납부’, ‘조회’, ‘등록’, ‘서비스’, ‘개인’, ‘홈페이지’, ‘소득세’, ‘모바일’, ‘발급’ 등의 단어들이 상위권에 위치하고 있다. 이러한 결과는 연결정도 중심성, 근접 중심성에서의 순위와 유사하다. 매개 역할에서는 홈페이지나 모바일의 개인 소득세 신고, 서비스 등록, 납부, 신청, 조회 등이 높다는 것을 나타낸다. 따라서 국세청입장에서는 개인 소득세 부문의 등록, 납부, 신청, 조회 탭들이 홈페이지와 모바일 앱 에서 유기적으로 접근할 수 있도록 세밀한 분석을 하여 업그레이드하는 것이 유용할 것이다.

또한 “홈택스” 화면의 조회/발급 탭의 내용은 발급위주로 구성되어 있다. 납세자가 신고/납부 탭에서 관심이 있는 세금에 대하여 얼마가 되는지, 절차가 어떻게 되는지 등에 대하여 사전적으

로 파악할 수 있도록 구성하면 더욱 사용자 관점의 접근이 될 것이다. 윤병섭·우상덕(2007)은 싱가포르의 경우에는 납세자가 사전에 적정과세 여부를 검토할 수 있도록 조치하고 있다고 하였다.

주요 단어들은 위세 중심성에서도 연결정도 중심성, 근접 중심성, 매개 중심성과 유사하게 위치하였다. 중복되는 단어들은 차지하는 중요성 비중도 높다는 것을 나타낸다. 특히 ‘신고’, ‘신청’, ‘등록’, ‘납부’, ‘조회’ 등의 단어들은 위세중심성 뿐만 아니라 모든 중심성에서 높은 순위에 존재한다. 사용자들이 “홈택스”에서 어떻게 신청하고 신고하며 납부하는가에 큰 비중을 두고 관심을 가지고 있다는 것이다. 사용자 입장에서는 익숙하지 않은 시스템이지만, 각 개인이 부담해야 하는 직접적인 세금이기 때문에 높은 관심을 가지고 있는 것이다. 따라서 사용자들이 “홈택스”에 접근 시에 쉽게 접근할 수 있도록 지속적인 개선이 필요할 것이다. 이러한 관점에서, 정성모(2015)는 화면과 메뉴 그리고 접근성이 좋은 직관적인 화면 구성의 개선을 주장하였다.

4. CONCOR 분석 결과

CONCOR 분석은 키워드 간 동시출현 매트릭스의 분석을 토대로 노드들의 블록을 형성하여 블록들 간의 관계를 파악하는 방법이다(안명숙과 오익근, 2015). 이 분석은 블록 유사집단을 군집으로 묶고 군집간의 연결 구조에 대한 관계성을 의미 있는 사회연결망을 분석하는 것이다. 분석도구는 유씨넷(Ucinet)과 넷드로우(NetDraw) 프로그램을 활용하였다.

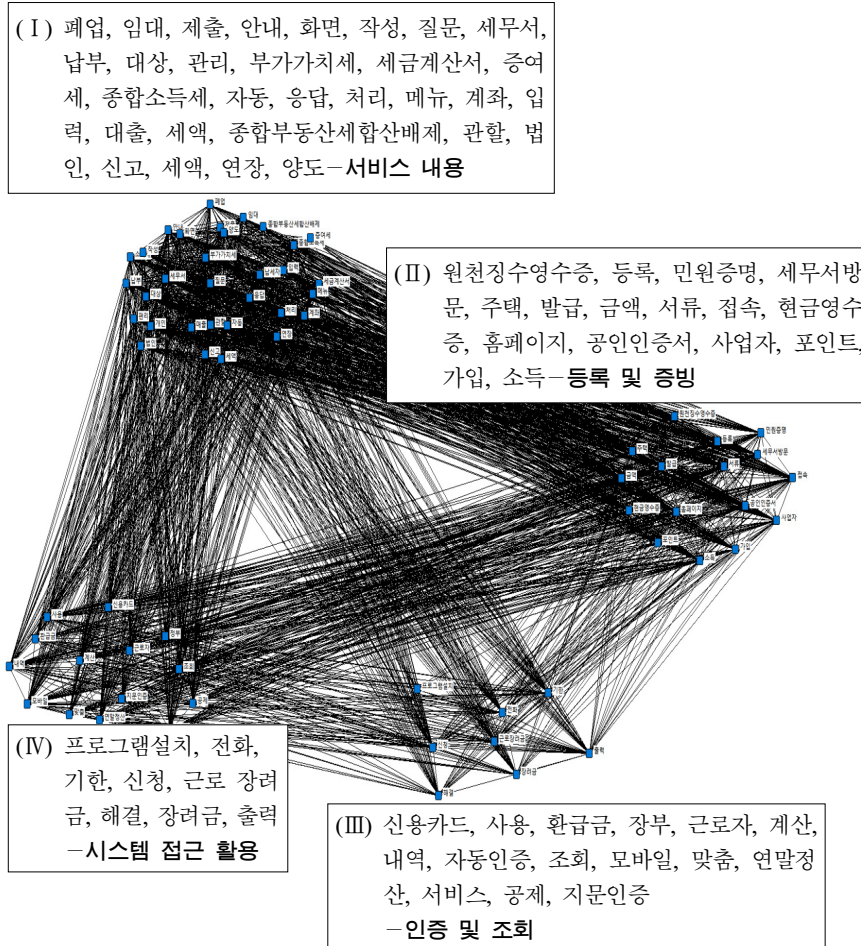
다음 <그림 2>에서 보는 바와 같이, 군집은 4개의 클러스터로 형성되어, (I)부터 (IV)군집으로 명명하였다. 각 군집들은 구성내용을 분석하여 그 특징에 대한 요약단어를 제시하였다. 그 내용은 다음과 같다.

첫째, (I)군집은 ‘폐업’, ‘임대’, ‘제출’, ‘안내’, ‘화면’, ‘작성’, ‘질문’, ‘세무서’, ‘납부’, ‘대상’, ‘관리’, ‘부가가치세’, ‘세금계산서’, ‘증여세’, ‘종합소득세’, ‘응답’, ‘처리’, ‘메뉴’, ‘계좌’, ‘입력’, ‘대출’, ‘자동’, ‘세액’, ‘종합부동산세합산배제’, ‘관할’, ‘법인’, ‘신고’, ‘세액’, ‘연장’, ‘양도’ 등의 단어들로 “홈택스” 내용으로 구성되어 있다. 즉 단어들은 부가가치세, 증여세, 종합소득세, 임대소득 등의 조세와 대상, 관할 세무서의 세액 및 안내에 대한 질문입력 및 응답, 안내 화면 및 메뉴관리, 자료 제출 작성, 자동 처리, 세액 대출, 계좌입력 등에 관한 관심내용으로 그룹을 형성하고 있다. 그리하여 (I)군집은 “서비스내용”들이 포함된 군집이라 할 수 있다.

둘째, (II)군집은 ‘원천징수영수증’, ‘등록’, ‘민원증명’, ‘세무서방문’, ‘주택’, ‘발급’, ‘금액’, ‘서류’, ‘접속’, ‘현금영수증’, ‘홈페이지’, ‘공인인증서’, ‘사업자’, ‘포인트’, ‘가입’, ‘소득’ 등의 단어들로 “등록 및 증빙”에 대한 내용으로 구성된 군집이라 할 수 있다.

셋째, (III)군집은 ‘신용카드’, ‘사용’, ‘환급금’, ‘장부’, ‘근로자’, ‘계산’, ‘내역’, ‘자동인증’, ‘조회’, ‘모바일’, ‘맞춤’, ‘연말정산’, ‘서비스’, ‘공제’, ‘지문인증’ 등으로 사용을 위한 “인증 및 조회”에 관한 관심사항으로 묶여 있다.

〈그림 2〉 “홈택스” 관련 데이터의 CONCOR 분석



넷째, (IV)군집은 ‘프로그램설치’, ‘전화’, ‘기한’, ‘신청’, ‘근로 장려금’, ‘해결’, ‘장려금’, ‘출력’ 등으로 “시스템 접근 활용”에 관한 내용으로 구성되어 있다.

(I)군집은 나머지 3개의 군집 중에서 (II)군집과 높은 상관관계를 유지하고 있는 것으로 시각적으로 질게 연결되어 2 군집간의 상관성이 가장 높게 나타나 있다. 즉 사용자들은 “홈택스”에서 부가가치세, 증여세, 종합소득세, 임대소득 등에 대하여 특히 관심을 가지고 있으며, 이에 따라 홈페이지 가입이나 직접 세무서를 방문하여 등록하는 방법, 소득에 대한 세금 금액과 그에 대한 증명서, 서류, 영수증, 포인트 등을 궁금해 한다는 것이다. 공급자인 국세청입장에서는 다수의 사용자들이 관심을 가지는 세목에 대한 궁금증과 예상세액 등에 대하여 좀 더 이해하기 쉬운 설명으로 업그레이드 하는 것이 유용할 것이다. 변정희 · 심태섭 · 김상현(2016) 등의 연구에

서는 “예상세액계산하기”가 사용자만족에 비유의적 영향을 미치는 것으로 나타났다. 사용자만족의 개선이 필요하다.

(I)군집은 (III)군집, (IV)군집과도 상관관계를 유지하고 있다. 즉 사용자들은 모바일 인증에 의한 사용방법과 개인별 신용카드, 장부에 의한 계산내역에 따르는 공제서비스, 환급혜택 등을 조회하여 공급증을 해결하는 맞춤형 정보에 관심을 가지고 있는 것이다. 그리고 이 군집은 개인별 관심 세금에 대한 프로그램설치에 따른 신청, 해결 및 출력, 전화해결, 장려금 등에 상관관계를 나타내고 있다. 그리하여 공급자인 국세청입장에서는 사업자 및 근로자의 주요 관심 세금 내역에 대한 세부적인 맞춤형 전화상담, 장려금 혜택, “홈택스” 사용의 쉬운 접근이 용이하도록 주력하는 것이 필요하다. 임재희 외(2007)의 연구에서도 “홈택스” 서비스 시스템 활용의 신뢰성은 이미 확보되었으나, 정보의 단순 입·출력뿐만 아니라 국세정보를 제공함으로써 적시성 있고 유용한 정보의 공급이 수반되어야 한다고 하였다.

(II)군집은 (III)군집과 (IV)군집, (IV)군집은 (III)군집과 상관관계가 높게 나타났다. 이것은 “홈택스” 접근과 설치, 인증 및 조회 등과의 관계성에 정도의 차이가 있지만 서로 연결되어 있다는 것을 나타낸다. 이 군집은 홈페이지 및 모바일, 전화, 세무서 방문에 의한 인증 접속에 의한 개인별 맞춤형 세금계산 해결, 근로자의 연말정산, 신용카드 납부조회, 계산내역, 영수증 발급 및 출력, 장려금 및 환급금 혜택 등에 상관관계를 형성하고 있다. 결국 사용자들은 관심 세금에 대한 즉각적이고 쉬운 프로그램 접근과 증빙에 대하여 관심을 갖고 있는 것이다. 심태섭과 송인국(2005)은 “홈택스” 서비스 시스템에서 시스템의 접근편리성, 제공정보의 품질, 문의사항에 대한 서비스, 시스템의 보안성, 제공정보의 최신성 등이 사용자 만족도에 유의적인 영향을 미친다고 하였다. 그러나 본 연구의 결과에서는 보안성 보다는 모바일, 자동인증 및 조회 등의 편리성과 사용자들이 우선적으로 원하는 내용 등으로 나타났다.

이와 같이 CONCOR 분석에서는 홈페이지와 모바일에서 “홈택스” 접근을 위한 편리성과 직접적인 세금관련 관심사항에 대한 설명과 쉬운 이해 등에 관심을 나타내고 있는 것으로 나타났다. 공급자인 국세청 입장에서는 사용자의 관심사항을 깊게 분석하여 개선해 나가는 것이 필요하다. 예를 들어, 연말정산 시즌에는 근로자들이 가장 관심을 가지는 세금정보를 분석하여 홈페이지 및 모바일에 시각적인 내용으로 확인할 수 있도록 일시적으로 탭을 구성하는 방법이 필요한 것이다. 이성욱 외(2015)의 연구에서는 IPA 분석결과, “홈택스”의 검색 용이성, 검색 조작의 간편성, 정보의 이해가능성에 대한 개선이 필요하다고 하였다. 또한, QFD의 분석결과에서는 내부자료 검색 엔진 및 디렉토리 서비스, 비공개 업무처리, 민원 업무 및 접수처리, 다운로드 속도, 질의응답에 대한 중요성을 강조하였다.

V. 결 론

2015년 국세청의 “홈택스” 시스템은 기존의 분리된 개별적 시스템을 하나로 통합하여 국세청 “홈택스”라는 새로운 프로그램을 실행하였다. 새로운 시스템은 사용자의 니즈에 맞는 수정·보완되기 위하여 다양한 분석이 실행될 필요가 있다. 하지만 통합된 “홈택스” 시스템에 관한 사용자 입장의 연구는 매우 미흡한 실정이다. 또한 국세청의 “홈택스”에는 사용자에게 대한 만족도 평가가 있으나 국세청의 인위적인 설문에 의한 사용자의 선택이다. 설문대상의 한계가 존재하는 것이다. 다수의 사용자가 스스로 의견을 제시한 내용을 조사하고 분석하여 개선하면 효과가 더 크게 나타날 것이다.

따라서 본 연구에서는 소셜 네트워크를 활용하는 수많은 “홈택스” 사용자들의 빅 데이터를 활용하여 무엇에 관심을 가지고 있는지를 분석하였다. 빅 데이터 분석에는 텍스트(Textom), 유씨넷(Ucinet), 넷드로우(NetDraw) 등의 프로그램을 활용하였다. 그 결과, 빈도분석에서는 ‘신청안내’, ‘연구’, ‘연말정산서비스안내’, ‘안내’ 등이 최상위 순위로 표시되었다. 따라서 국세청은 “홈택스”에 관한 기존의 신청과 안내, 연말정산서비스안내 등을 세밀하게 연구하여 소비자 관점의 개선을 시도하는 것이 유용할 것이다. 예를 들어 “홈택스”의 연말정산서비스 탭에서는 연말정산에 대한 개략적인 예상세액 탭이 존재한다. 그러나 일반인들은 각 세목에 대하여 익숙하지 못하다. 연말정산 시즌에는 일반인들이 쉽게 이해할 수 있는 내용의 풀이와 예제, 소득공제가 아닌 세액공제 영향 등을 참고 팝업으로 제공하는 것도 하나의 대안이 될 수 있을 것이다.

또한 주요단어에는 ‘사업자등록’, ‘영향’, ‘원천징수’, ‘관할’, ‘세금’, ‘개선방안’, ‘쌍커피’, ‘근로장려금’, ‘장려금’, ‘신고’, ‘확정일’, ‘해외금융계좌’, ‘적극힌트’, ‘경우’, ‘귀속’, ‘사이버시스템’, ‘납세자’, ‘홍보대사’, ‘방법’, ‘세금상식’, ‘근로소득’, ‘공정위’, ‘납부’, ‘소득세’, ‘편리’, ‘지속가능’ 등이 추출되었다. 이러한 결과는 사용자들의 개별적인 맞춤내용에 대한 초점이 나타난다. 즉 선행연구에서는 신뢰성과 보안성 등의 시스템적인 측면에서 분석한 비중이 높다고 하였다. 그러나 본 연구에서는 향후에는 사용자들이 관심을 가지고 있는 주요 개별적인 단어에 대한 내용들의 서비스에 초점을 맞추어 시스템을 구성하는 것이 유용하다는 결과로 나타났다.

전체네트워크 분석에서는 ‘신고’, ‘신청’, ‘금액’, ‘서비스’, ‘발급’, ‘사업자’, ‘조회’, ‘납부’, ‘접속’ 등이 중심에 있다. 이것은 국세청 “홈택스” 사이트에서 사업자가 어떻게 신고, 신청, 조회, 발급, 납부를 하는지에 대한 접속방법과 금액이 얼마인지 등에 대한 사항에 집중적인 관심을 가지고 있다는 것을 나타낸다. “홈택스” 사용자들은 계속적으로 시스템의 사용방법에 대하여 집중적인 관심을 가지고 있다. 따라서 보다 적극적으로 쉽게 활용할 수 있도록 하는 방법을 강구할 필요가 있다.

주요 단어들 간의 연결성을 나타내는 중심성 분석에서는 빈도분석과 유사하였다. 추가적으로는 ‘홈페이지’, ‘모바일’, ‘가입’, ‘편리’ 등에 연결정도 중심성이 높은 순위에 존재하여 편리한

홈페이지와 모바일 가입의 연결된 유용성에 관심이 존재하였다. 또한 근접 중심성에는 ‘신고’, ‘납부’, ‘신청’, ‘등록’, ‘서비스’, ‘조회’, ‘개인’, ‘홈페이지’, ‘모바일’ 등이 존재하였다. 사용자들은 홈페이지나 모바일에 의해서 “홈택스” 등록 및 신청, 신고 및 납부·조회 등을 시스템에서 동시에 확인 이슈에 관심이 높다는 것이다. 국세청은 사용자 입장에서 편리하게 접근이 될 수 있도록 홈페이지와 모바일에 대하여 동일한 탭을 형성하여 익숙하게 활용될 수 있도록 노력하여야 할 것이다.

본 연구에서는 다른 선행연구와는 다르게, 매개 중심성과 위세 중심성이 연결 중심성, 근접 중심성에서의 순위와 유사하였다. 이러한 결과는 중복되는 단어들의 비중이 높다는 것을 나타낸다. 즉 쉽고 편리한 “홈택스” 시스템의 구성에 집중할 필요가 있다는 것이다. 하지만 위세 중심성에서 ‘홈페이지’, ‘모바일’, ‘소득세’, ‘신용카드’, ‘편리’, ‘가입’ 등의 단어들이 타 중심성에 비하여 상대적으로 높은 관심 순위로 나타났다. “홈택스”에서는 홈페이지와 모바일에 대한 활용을 제공하고 있다. 그러나 사용자관점에서의 세밀한 개선이 필요한 것이다. 예를 들어 홈페이지와 모바일의 동일하거나 유사한 화면 구성을 함으로서 익숙한 접근이 가능하도록 하는 것이 유용할 것이다. 또한 사용자들은 신용카드에 대한 관심도가 매우 높기 때문에 사용자관점의 세밀한 요구사항을 파악할 필요가 있다.

군집 간의 의미 있는 사회연결망을 도출하기 위한 CONCOR 분석에서는 (I)군집은 (II)군집과 가장 높게 상관성을 가지고 있었다. 즉 “홈택스” 사용자들은 홈페이지나 세무서 방문 등록을 한 후에는 소득세 납부 금액과 증명서, 서류, 영수증, 포인트 등을 개별적으로 확인하는데 관심을 가지고 있기 때문에 가장 우선적인 접근이 가능하도록 하는 것이 만족도를 높이게 되는 것이다. (I)군집은 (III)군집, (IV)군집과도 상관관계를 갖고 있다. 또한 (II)군집은 (IV)군집, (II)군집, (IV)군집은 (III)군집과 상관관계가 높게 나타났다. “홈택스” 사용자들은 홈페이지 및 모바일 인증에 의하여 세금의 계산내역과 장려금 공제, 환급, 신용카드 납부에 관계성을 나타내고 있다. 그리고 납부에 의한 상관성은 전화 상담에 의한 문제해결에서도 나타난다. 개인별 맞춤을 위한 전화 상담이 편리하고 신속한 조치가 될 수 있도록 하는 방안이 필요하다.

본 연구에서는 새롭게 통합된 “홈택스”에 관한 사용자 관심이슈를 빅 데이터에 의하여 도출하였다. 기존의 실증분석이나 사례연구에서는 연구대상의 양적 한계가 존재하였다. 본 연구에서는 소셜 네트워크상의 빅 데이터를 활용함으로써 실질적 결과에 의한 일반화에 더욱 접근하게 된 것이다. 연구결과는 현실적인 데이터를 분석하였기 때문에 공급자인 회사의 입장에서 사용자의 니즈(needs)를 파악하는데 실질적인 유용한 자료가 될 것이며, 사용자 만족을 상승시키기 위한 대응책으로 실행해야 할 직접적인 내용으로 표출되는 것이다(황옥선, 2017).

향후에는 통합된 “홈택스” 시스템 내에서 각 모듈에 대한 조세 세목별, 기간별 분류에 의한 구체적인 해결방안을 분석하는 연구도 필요하다. 또한 2015년 통합되기 전과 통합 후의 비교분석, 통합 후의 긴 기간 동안의 변화에 대한 분석은 향상된 개선방향을 찾는 도출의 연구로 유용성을 갖게 될 것이다.

참고문헌

- 국세청. 2005. “세계 최고 수준의 전자세정 구현 - “홈택스” 서비스 시행 3년간의 운영성과-”. 국세청 보도 참고자료 : 1-4.
- 김명종 · 송민정. 2013. “홈택스” 시스템의 품질요인이 사용자 만족도에 미치는 영향”. 세무와회계저널 제14권 제6호 : 133-162.
- 김현정 · 조남옥 · 신경식. 2015. “항공산업 미래유망분야 선정을 위한 텍스트 마이닝 기반의 트렌드 분석”. 지능정보연구 제21권 제1호 : 65-82.
- 기획재정부. 2011. “국세통합시스템[TIS] 전면개편 사업타당성 확보”.
- 노규성 외. 2019. “경영정보시스템”. 광문각.
- 배춘호. 2009. 국세청 ““홈택스”서비스가 납세순응에 미치는 영향에 관한 연구”. 강남대학교 박사학위논문.
- 배춘호 · 서희열 · 배수진. 2011. “세정 “홈택스” 서비스 시스템이 납세순응에 미치는 영향에 관한 연구”. 세무학연구 제28권 제1호 : 225-367.
- 변정희 · 심태섭 · 김상현. 2016. “국세청 “홈택스” 연말정산간소화와 편리한 연말정산 서비스에 대한 사용자만족도 및 성과에 관한 연구”. 세무학연구 제33권 제4호 : 43-85.
- 서정원 · 김형중. 2019. “자금세탁 의심거래 탐지 방안 연구-소셜 네트워크 분석 기법을 중심으로”. *Journal of Digital Contents Society* 20(3) : 569-576.
- 심태섭 · 송인국. 2005. “국세청 “홈택스”서비스 시스템에 대한 사용자만족도와 업무효율향상 연구”. 세무학연구 제22권 제4호 : 41-62.
- 아마르자르갈 다그와더르지 · 이여라 · 김학선. 2020. “빅데이터 분석을 활용한 베이커리에 대한 소비자인식에 관한 연구”. *Culinary Science&Hospitality Research* 제26권 제4호 : 66-75.
- 안명숙 · 민용기. 2016. “5성급 호텔 이용에 대한 동기 빅 데이터 분석 : 동기이론 Push-Pull factor를 중심으로”. 관광연구 제31호 제5권 : 87-104.
- 안명숙 · 오익근. 2015. “네트워크 텍스트 분석 기법을 적용한 특급호텔 패키지 이용에 대한 태도 분석 -포털사이트 활용-”. 관광연구 제30호 제5권 : 163-181.
- 오익근 · 이태숙 · 전채남 · 윤영일. 2015. “빅 데이터 분석을 통한 한국관광 인식에 관한 연구”. (사)한국관광학회 국제학술대회 : 595-602.
- 이성욱 · 오상훈 · 김현수. 2015. “QFD를 활용한 국세청 “홈택스” 시스템 품질 제고에 관한 연구”. 대한경영학회지 제28권 제2호 : 505-528.
- 이수진 · 전유나. 2016. “소셜미디어 빅데이터 분석을 통한 관광 인식조사 -경기도 지역을 중심으로-”. GRI연구논총 제18권 제1호 : 83-109.
- 21세기 정치학대사전. 2017. “네트워크 분석”.
- 임재희 · 신성식 · 양해면. 2007. “HTS 시스템품질변수가 HTS성과에 미치는 영향에 관한 연구”. 국제회계연구 제18권 : 301-320.
- 장아름 · 박지영 · 박한우. 2015. “트위터 데이터를 활용한 부산국제영화제(BIFF)의 글로벌 커뮤니케이션 전략 연구”.

- 니케이션 네트워크 분석”. 한국자료분석학회지 제17권 제2호 : 939-949.
- 정명재 · 노학래 · 박민식. 2014. “SNS 빅 데이터를 활용한 철도관광 이용객들의 반응 조사”. 한국철도학회 학술발표대회 논문집 : 456-461.
- 정성모. 2015. “통합된 “홈택스” 서비스 시스템에 대한 사용자만족도와 성과에 관한 연구 -세무대리인을 중심으로-”. 단국대학교 박사학위논문.
- 정성모 · 변정희 · 김상현. 2017. “통합된 “홈택스” 서비스 시스템에 대한 사용자 만족도와 성과에 관한 연구”. 회계 · 세무와 감사 연구 제59권 제2호 : 201-226.
- 채혜린. 2018. “공공기관 앱 사용자수 1위는 코레일 톡, 국세청 “홈택스” 사용자는 69만명 4위”. NTN.
- 최동춘 · 박종원 · 차현수. 2014. ““홈택스” 서비스(Home Tax Service)의 재이용의도에 관한 연구”. 재무와회계정보저널 제14권 제4호 : 75-93.
- 한경훈. 2011. “국세청 “홈택스”시스템 수용에 영향을 미치는 요인에 관한 연구”. 국제회계연구 제35집 : 315-332.
- 황옥선. 2005. “전자세무신고 “홈택스”(HTS)의 유용성에 관한 연구”. 세무와회계저널 제6권 제3호 : 83-97.
- 황옥선 · 민정기. 2017. “빅 데이터 활용에 의한 리조트 관광객의 소비성향 분석”. 호텔리조트학회지 제16권 제2호 : 5-26.
- 황옥선. 2017. “빅 데이터에 의한 해상공관의 관심 트렌드 분석”. 호텔리조트학회지 제16권 제2호 : 125-145.
- 황옥선. 2020. “빅 데이터 활용에 의한 상업교육 관심 트렌드 분석”. 상업교육연구 제34권 제1호 : 1-22.
- Baggio, R. and C. Cooper. 2010. Knowledge transfer in a tourism destination : the effects of a network structure. *The Service Industries Journal* 30(10) : 1757-1771.
- Douglas, L. 2020. The importance of Big Data : A definition Gartner. Retrieved 21 June.
- Spencer A. Wood, Anne D. Guerry, Jessica M. Silver and Martin Lacayo. 2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports*. 3.
- Firat Tekiner and A. John Keane. 2013. Big Data Framwork. 2013 IEEE International Conference on Systems, Mans, and Cybermetics : 1494-1499.
- Cho, J. S. 2012. Inflow and outflow analysis of double majors using social network analysis. *Journal of the Korean Data & Information Science Society* 23 : 693-701.
- Choi, S. B., Kang, C. W., Choi, H. J. and Kang, B. W. 2011. Social network analysis for a soccer game. *Journal of the Korean Data & Information Science Society* 22 : 1053-1063.
- Stephen P. Borgatti. 2005. Centrality and network flow. *Social Networks* 27 : 55-71.
- Wasserman, S., and K. Faust. 1994. *Social Network Analysis : Methods and Applications*. Cambridge : Cambridge University Press. 8 : 215-220.

Analysis of Home Tax Utilization by Extraction of Social Network Big Data Interest Issues

UgSun Hwang*

〈Abstract〉

Hometax service in 2015 is an integrated system of a new concept, and various analyzes are needed for modification and supplementation to satisfy user needs. Therefore, in this study, we analyzed what interests in the system by utilizing the big data of numerous hometax service users using social networks. For Big Data analysis, programs such as Textom, Ucinet, and NetDraw were used.

As a result, in frequency analysis, ‘application guide’, ‘research’, ‘year-end settlement service guide’, and ‘guide’ were ranked in the top ranking. Interest in year-end settlement guidance, application, and research was high. In the overall network analysis, ‘report’, ‘application’, ‘amount’, ‘service’, ‘issuance’, ‘operator’, ‘inquiry’, ‘payment’, and ‘access’ are at the center. This means that the “Home Tax” site is focused on the access method and the actual tax amount for how business operators report, apply, inquire, issue, and pay.

In centrality analysis and CONCOR analysis, there existed connectivity between clusters in hometax registration and application, report and payment, inquiry, income tax payment amount and certificate, documents, receipts, points, and telephone consultations by homepage or mobile. The IRS needs to take action on these words.

〈Key words〉 Home Tax, Big Data, centrality analysis, CONCOR analysis, interests

* Professor, School of Global Business, Halla University, yesek41@naver.com