

빅데이터에 기반한 지역 상점 관련 정보제공 서비스

Big data-based Local Store Information Providing Service

문창배*, 박현석**

한양사이버대학교 전기전자통신공학부*, 한양대학교 정보시스템학과**

Chang-Bae Mun(changbae@hycu.ac.kr)*, Hyun-Seok Park(hp@hanyang.ac.kr)**

요약

빅데이터를 활용한 위치정보 서비스가 지속적으로 발전하고 있다. 네비게이션의 측면에서는 지도 API서비스부터 선박 항해정보에 이르기까지 서비스의 범위가 확대되었고, 시스템 응용정보로서는 각 위치에 대한 SNS와 블로그 검색 기록에까지 확장되고 있다. 특히 최근에는 위치기반 검색 및 광고, 무인 자동차, IoT(Internet of Things) 및 O2O (Online to Offline) 서비스 등 신규 산업으로 활용되고 있다. 본 연구에서는 사용자가 특정 경로를 이동할 때, 빅데이터를 활용하여 근처 상점에 대한 정보를 보다 효과적으로 수신할 수 있도록 하는 인터넷 서비스를 제안하였다. 또한 지역 상점은 이 시스템을 사용하여 저렴한 비용으로 효과적으로 홍보할 수 있도록 시스템을 설계하였다. 특히 실시간으로 웹 기반 정보를 분석하여 사용자에게 제공하는 상점 정보의 정확도를 높이고자 하였다. 이 시스템을 통해 시스템 이용자인 일반 사용자와 상점 업주는 보다 효과적으로 정보를 활용할 수 있게 될 것이다. 또한 시스템 관점에서, 다양한 웹서비스와 통합하여 신규 서비스를 창출하는 것에 활용할 수 있을 것이다.

■ 중심어 : | 위치기반서비스 | LBS 서비스 | 아파치 스파크 | 빅데이터 | 추천 서비스 |

Abstract

Location information service using big data is continuously developing. In terms of navigation, the range of services from map API service to ship navigation information has been expanded, and system application information has been extended to SNS and blog search records for each location. Recently, it is being used as a new industry such as location-based search and advertisement, driverless cars, Internet of Things (IoT) and online to offline (O2O) services. In this study, we propose an information system that enables users to receive information about nearby stores more effectively by using big data when a user moves a specific route. In addition, we have designed this system so that local stores can use this system to effectively promote it at low cost. In particular, we analyzed web-based information in real time to improve the accuracy of information provided to users by complementing the data. Through this system, system users will be able to utilize the information more effectively. Also, from a system perspective, it can be used to create new services by integrating with various web services.

■ keyword : | Location based Service | LBS | Apache Spark | Big Data | Recommendation Service |

* 본 연구는 한양대학교 교내연구지원사업으로 연구되었음(HY-2018)

접수일자 : 2020년 01월 22일

수정일자 : 2020년 02월 05일

심사완료일 : 2020년 02월 05일

교신저자 : 박현석, e-mail : hp@hanyang.ac.kr

I. 서론

Technavio의 Location-Based Service (LBS) 사업 관련 보고서에 따르면, 국제 LBS 시장은 2020년까지 198억 달러까지 발전할 전망이다. 특히 '위치기반의 검색과 광고(35억 달러)'가 가장 큰 비중이고 무인자동차, Internet Of Things (IOT), 자동 지역광고 등의 신규 사업 자원으로 분석되고 있다[1]. 스마트폰에서 수집한 위치정보를 기반으로 한 LBS 사업의 초창기에는 다양한 위치정보 Application Programming Interface (API)로부터 위치의 특성을 분석하여 개별 어플리케이션으로 제공하는 서비스로서 발전되었으며 실내 측위, 자동차 및 선박의 네비게이션에 활용되다가 점점 그 범위를 확장하여 광고와 위치기반 마케팅, 위치기반 AR게임, 드론을 도입한 서비스 등에 활용되고 있다. 공공 서비스 측면에서는 사회안전망으로서 구조가 필요한 상황에서 구조작업을 위해 위치정보를 조회하는 기술도 점점 발전하고 있다. 이렇게 개별적인 위치정보를 활용하는 사례는 발전하고 있지만, 기존에 제시된 다양한 위치정보를 조합하고 정규화 과정을 수행하여 보다 정확한 정보로 재가공하는 기술에 대해서는 여전히 개발이 필요한 실정이다. 한국인터넷진흥원은 LBS의 서비스와 콘텐츠에 대해 위치정보 제공서비스와 위치정보 활용 서비스로 구분하여 LBS 산업의 범위를 제시하고 있다[2]. 이는 단순한 위치정보의 제공과 위치정보의 활용을 통해 보안, 지도 주변정보, 관제, 지역생활정보 등을 별도의 범주로 다룸으로서 위치정보의 확장성을 보여주는 것으로 볼 수 있다. 본 연구에서는 지도 API, SNS, 인터넷 검색정보, 인터넷 등록자료를 실시간으로 분석하여 특정 지역에 대한 복합적인 정보를 일목요연하게 정제하여 사용자의 요구에 맞는 정보를 제공할 수 있는 기술을 모색해본다. 본 논문은 다음 장에서는 관련 연구자료를 분석하여 현재 위치기반 시스템의 현상을 다루고 3장에서는 빅데이터 분석을 활용한 정보의 확장성과 정제 방안을 다룬다. 4장에서는 본 시스템의 결과에 따라 효용성을 확인하고 마지막 장에서는 연구의 결과에 대해 논평한다.

II. 본론

1. 신규 시스템의 필수 기능 분석

1.1 LBS 서비스의 발전

위치정보서비스는 세계적으로 2014년 이후 연평균 20%이상 성장을 거듭하여 2020년에는 43조 5천억원에 이를 것으로 예상되고 있다. 특히 위치정보와 관련한 빅데이터 분석, O2O 서비스 및 IOT서비스, 무인자동차 기술 등이 활성화되어 위치 정보 사업이 4차 산업혁명시대에 도래할 융합 산업 발전의 중심이 될 것으로 전망되고 있다. 스마트폰의 도입 초기에는 지도정보 API이외의 수익모델에 대해 부정적인 인식이 있었으나 점점 지도 주변정보와 안전서비스, 엔터테인먼트의 다양한 분야에서 새로운 서비스가 출현하며 융합영역의 LBS 산업이 확대되고 있다[3]. 이렇게 위치정보사업의 시장은 발전가능성이 크지만, 국내에서는 규제 기반의 법률과 정책이 부재하여 LBS사업의 활성화가 저해되고 있는 실정이다. 이에 한국인터넷진흥원에서는 위치정보에 대한 규제혁신을 통한 산업 활성화 방안 발표하여, 지원 체계 목표도 2020을 수립하여 법과 제도적인 개선을 통해서 신규 LBS사업의 준비와 더불어 사업의 도약기까지 지원하는 체계를 정비하고 있다. 이를 통해 서비스의 출시부터 네트워킹, 홍보, 마케팅과 같은 서비스 확대의 측면까지 지원함으로써 신규 LBS 스타트업의 발전할 수 있는 발판 마련의 준거가 될 수 있다[4]. 또한 네이버는 에어스페이스(AiRSPACE) 서비스에서 인공지능에 기초하여 사용자의 현재 위치와 사용자의 관심사에 맞춘 주변 장소를 추천해주는 서비스를 마련하였다. 이 서비스는 위치, 시간, 나이등의 정보에 따라 맞춤형 위치 정보를 제공하는 것으로 추천시스템을 더욱 고도화하여 사용자에게 적합한 장소를 추천하도록 하였다[5]. 중고거래 플랫폼인 당근마켓은 근거리 위치정보를 최우선 중심 정보로 삼고 있으며, 우선 사용자의 위치를 인증하고 인근 지역의 사용자와 직거래를 지원하는 기능을 제공한다[6]. 특히 당근마켓에서는 일상적인 자신의 동네에 관한 정보제공이라는 측면에서 일상 활동에 대한 커뮤니티를 제공하고 이 구매활동은 하나의 어플리케이션을 활용한 활동으로서 직거래 구매 행위를 지원하는 요소로 작용한다. 이렇게 위치정보가

타 정보와 함께 융합되어 결합된 양상의 새로운 정보를 창출하는 요소로 진화되고 있음을 알 수 있다.

1.2 융합 위치 정보의 마케팅에 활용

온라인과 모바일 결제의 비중이 점점 증가하고 있음에 따라 마케팅의 대상도 이동해갔지만, 오프라인 매장은 고객과 브랜드가 직접적으로 대면하는 주요 접점으로서 작용하고 있다. 특히 브랜드에 대한 고객 경험의 측면에서 소비자와의 소통을 중요시하는 각 제품의 특성에 따라 중요한 요소가 된다. 이에 저전력 블루투스를 활용한 비콘 마케팅을 도입하여 스마트폰 어플리케이션으로 할인 쿠폰이나 프로모션 정보를 비콘으로 전달하는 방식이 많이 사용되고 있다. 이는 매장 인근의 고객들에게 프로모션 정보와 제품을 직접적으로 홍보할 수 있는 수단으로서 매출에 기여할 수 있고 고객 입장에서든 인근에서 편리하게 정보를 제공받을 수 있으므로 편리한 이점을 가진다. 또한 인터넷 몰의 구매 데이터와 접속로그를 기반으로 사용자의 구매정보와 지역정보, 날씨 등의 정보를 종합하여, 고객행동 특성에 기반한 마케팅과 프로모션도 진행할 수 있으며 오프라인 매장 내에서의 구매자의 동선의 패턴을 분석하여 이를 온라인의 메뉴 설정에 활용할 수도 있고 이전에 방문한 매장 정보에 따른 고객 단위의 타겟 마케팅을 수행할 수도 있다. 기존에는 CCTV상의 고객 이동데이터 분석이나 실제 시뮬레이션 대상 고객의 설문에 의해서만 가능하였던 정보이지만, 이제는 고객이 오프라인 몰의 스마트폰 어플리케이션에서 위치정보 제공에 동의한 경우에 위치신호를 분석하여 주로 머문 매장과 경로를 분석하여 분석할 수 있으며 이는 미래 유통산업의 주요 분석 데이터로 활용될 수 있을 것이다. 예를 들어 백화점의 식당가에서는 음식점의 유입인원 분석을 통하여 외부 인테리어와 외부 메뉴판, 매장 앞의 체류시간, 재방문, 테이블 배치, 날씨 등이 실제로 매장의 매출에 영향을 미치는 정도를 보다 정확히 분석가능하게 되는 것이다. 또한 마트에서는 이와 같이 고객의 동선을 분석하여 고객 별 구매패턴을 분석하여 상품 배치 플랜을 재수립하고 직원의 배치와 응대패턴, 매장 배치에 대해서도 보다 주요한 인사이트를 도출하는 것에 활용될 수 있을 것이다.

2. 선행연구에 대한 분석

2.1 위치정보의 융합적인 활용

이윤주의 연구에서는 국내 관광업계의 서비스 제고를 위한 모바일 위치정보 통합 시스템을 제안하였고 여행사에서 단체 관광객들을 안정하게 통솔할 수 있는 정보시스템을 제시하였다[7]. 한국마케팅연구원은 위치기반 서비스의 마케팅을 통해 성공적으로 매출로 이어지게 하기 위해서는 소비자가 참여할 수 있는 방식을 사용해야 함을 제시한다. 즉 소비자에게 직접 자신이 참여한 것에 대한 경험을 주고 이를 통해 상품에 대한 관여정도를 높이고 실제 상품에 대한 구매 욕구를 상승시키는 과정에 대해 제시하였다[8]. 손동진의 연구에서는 인플루언서를 통해 소비자 지향적인 콘텐츠를 기획하고 관련 이벤트와 캠페인을 제시하여 성과에 대한 전략을 제시하였다[9]. 이를 위치정보와 조합하면 특정 지역의 유력 정보제공자를 통해 특정 마케팅 대상에 대해 소비자의 구매결정에 영향을 끼칠 수 있는 요소를 찾고 이를 유튜브 유명인의 유튜브 방송과 같은 지역적 콘텐츠 제작에 활용하여 마케팅 효과를 높이는데 이용하는 것으로 활용할 수 있다. 현재 월 거래액 6000억에 이르는 배달 어플리케이션과 같이 지역 음식점 배달정보를 기반으로 마케팅과도 연계할 수 있으며 향후 자율주행 차량이나 드론과 같은 이동장치에 광고 메시지를 노출하여 특정 지역을 기반으로 홍보에 활용할 수도 있다. 또한 공연장에서는 복수의 입구 단위의 혼잡도를 분석하여 도출된 동선을 바탕으로 보다 효율성이 높은 위치에 가판을 배치할 수 있을 것이며, 전시장에서는 고객의 이동 동선을 분석하여 보다 효과가 높은 위치에 인지도 높은 브랜드를 배치하거나 전시물의 배치를 수정하는 등의 최적화 작업을 시행할 수 있게 될 것이다. 고속도로 휴게소와 같은 지속적인 외부 고객이 유입되는 위치에서는 상,하행선에서 시간과 날씨, 고객의 최종 목적지에 따라 데이터를 분석하여 판매전략을 수립할 수 있다. 또한 사회안전망 차원에서는 시각장애인인 이동성에 위치정보를 빅데이터와 연계하여 제공하는 연구가 수행되었으며 이는 VR과 연계하여 보다 안정성을 높이는 결과로 제공될 수 있다[10][11].

2.2 빅데이터 위치정보 시스템

하둡(Hadoop)과 같은 빅데이터 분산처리 시스템을 통하여 디스크의 입출력 분산과 네트워크 입력과 출력에 대한 부하 분산이 수행될 수 있으며 데이터 리밸런싱을 통하여 데이터 노드에 대한 재배치를 수행하여 스토리지의 안정적 확장과 시스템적 적용성이 높아지는 효과가 있다. 네트워크 토폴로지를 기반으로 하여 데이터의 안정성을 확보할 수 있고, 메타 데이터에 대해서 네임노드에서 중앙 집중적 분석 과정을 제공함으로써 데이터 가용성을 확보하고, 데이터 노드에 이슈가 발생하더라도 데이터 손실 문제를 최소화할 수 있게 된다. 하지만 네임 노드의 이중화 문제, 파일 읽기가 급증하는 서비스에서의 이슈 등이 현재 상용 분산처리 서비스에서 해결 과정에 있다. 분산처리 서비스에서 스토리지 서버의 기능을 넘어서 최적화된 분석 처리와 연관된 맵리듀스를 적용함으로써 최고의 성능적용이 가능하다. 특히 하둡 맵리듀스는 HDFS에 인용된 대형의 데이터를 분석하기 위한 기법중의 하나로서 맵 태스크 혹은 리듀스 태스크 중의 태스크는 서로간의 중첩이 없으며 자신만의 데이터를 분석하고 처리하므로 중첩이 없는 병렬처리 시스템을 제공한다. 또한 블록단위로 데이터가 처리되고 복제 및 저장되므로 동일한 형태의 키값의 도식으로 데이터 형태를 관리하므로 데이터의 일관성을 유지하고 맵리듀스 과정에서의 최상의 결과를 보장하게 된다. 이와 같은 빅데이터 분산처리 시스템을 통해 특정 경로 이동자의 움직임을 분석함으로써 최적의 경로를 분석하는 연구가 수행되었다[12]. 한편 김은백의 연구에서는 대표적인 SNS인 인스타그램의 정보를 크롤링으로 분석하여 지역 활성화를 분석하였다[13]. 김태현의 연구는 스마트카드의 데이터를 이용하여 이동패턴에 대한 분석방안을 제시하였다. 이와 같은 기법들을 조합하면 현재 다양한 SNS 플랫폼을 크롤링 서버에서 분석하여 추론 연구에 활용하는 기법을 LBS 서비스 연구에 도입한 것으로 융합적인 측면의 새로운 용도를 제안한 것으로 볼 수 있다.

III. 시스템 분석, 설계, 구현

1. 시스템 분석과 설계

1.1 시스템의 구조 모델 수립

전체 본 시스템은 전체적으로 크게 세 가지의 기능으로 구성된다. 첫 번째 기능은 데이터 수집의 측면으로 특정 경로를 이동하는 사용자의 정보를 중앙서버에서 수집하고 분석을 수행한다. 그리하여 해당 경로에서 사용자가 관심을 가질 수 있는 특정 상점 정보를 찾고 유관한 상점의 마케팅 정보가 존재하면 그 정보를 사용자가 기존에 선택한 방식으로 푸시, 문자, 알림전화의 방식으로 제공해준다. 이를 위해 사용자의 기존 6개월간의 이동정보와 이동시의 방문한 상점의 정보, SNS, 날씨 데이터를 수집하여 분석하도록 한다. 그리고 사용자의 취향을 분석하여 인접도가 높은 사용자와 유사한 정보를 갖는 사용자의 정보와 비교하여 해당 사용자가 방문한 상점에 대해 가중치를 부여하도록 한다. 따라서 다양한 상점이 다수 존재할 때, 사용자의 이동방향에 따라 어떤 상점에 방문할 확률이 가장 높을지를 연산하여 최우선 상점부터 고객에게 추천과 알림을 시행한다. 이 전체적 구성에 대해서는 [그림 1]에 도식화 하였다. 두 번째 기능은 추천 결과에 대한 예외처리와 정보의 재구성이다. 본 시스템을 다수의 사용자가 동시에 이용한다고 가정할 때, 실제 고객 성향으로서는 방문한 것으로 결과가 나왔어야하나, 정반대의 성향을 갖는 고객이 방문하게 되는 예외상황이 나타날 수 있다. 이 케이스에서는 해당 고객의 방문정보들에 대해 해당 상점에 정보요청을 실시하여 예외상황에 대한 분석 자료를 수집하도록 한다.

따라서 테스트 지역을 넘어서는 범위에서 유사한 케이스가 발생할 경우에도 정확한 데이터 분석이 가능하

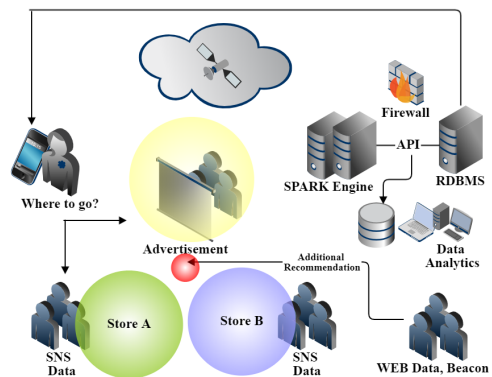


그림 1. 시스템 구조 모델링

도록 한다.

세 번째로는 상점이 상점의 주요 마케팅 정보를 중앙 서버에 업로드하고 동시에 개선된 정보를 지속적으로 피드백 받는 기능이다. 단순히 상점의 정보를 업로드하는 것이 아니라 상점을 이용한 고객 정보를 시스템에서 피드백 받고 마케팅 전략 수립에 보정을 할 수 있도록 도와주는 기능이다.

이를 통해 단순히 홍보수단으로서 시스템의 한계를 넘어서 미래에 상품 홍보의 방안을 결정하는 중요한 기반정보로 사용할 수 있게 된다.

정보제공 주체로부터 정보를 수집 및 분석처리를 위한 아파치 스파크(Spark)를 도입하고 분산전달을 위한 카프카(KAFKA), 하둡을 분석 모듈에 적용한다. 이 구조에 추가적인 작업을 위한 FLUME 모델을 이용하여 시스템 구조를 구현하였다. 중앙서버에서는 사용자의 정보와 개별 위치를 2초 단위로 수집하여 지속적으로 서버에 적재하도록 하였으며 중앙서버에서는 사용자의 성향과 처리 결과를 실시간으로 분석하여 현재 이후의 데이터 분석에 현재 데이터를 사용할 수 있도록 준비한다. 이렇게 기존의 정보들이 계속 빅데이터 처리서버에 축적되어 그 정보들이 하나의 기존 노하우로 작용하여 최종 사용자에게 최적의 정보를 제공할 수 있도록 한다. 이와 같은 전체 프로세스를 진화적 정보시스템으로 설계하기 위해 처리 속도의 이슈를 해결하기 위해 실시간 분산처리 시스템을 도입하였고, 내부와 외부 사용자에게 HTTPS를 통한 편리한 정보 제공을 위하여 Spring 프레임워크에 기초한 Restful API서버와 관계형 데이터베이스를 설계하였다. 이 빅데이터 분산처리 시스템에서 시간단위 DB배치작업으로 정제되고 분석된 정보들 중 빈도가 높거나 이동자의 가장 근거리해당하는 정보들을 우선 제공하고 그 정보의 결과를 저장하는 구조로 설계하였다.

이와 같은 정보처리 프로세스를 안정적으로 구현하기 위해 빅데이터 분산처리 환경을 도입하였다. 특히 안정적으로 내부와 외부의 사용자에게 연동정보를 제공하기 위하여 Spring framework로 구현한 Restful API서버, SOAP를 도입하였다. 특히 빅데이터 분산처리 시스템에서 동일한 배치작업으로 처리될 정보들 중 유사도나 사용빈도가 높은 정보들이 시스템에서 우선

처리 및 저장되는 구조로 구축하였다.

1.2 시스템의 전체 프로세스 설계

본 연구에서 제안하는 시스템의 Flume모듈은 오픈 소스로서 로그 컬렉션을 제공하며 시스템의 서버 로그, 정형, 비정형 데이터를 대상으로 데이터 수집을 시행한다. 다양한 데이터 포맷과 다양한 네트워킹 프로토콜을 지원하며 별도의 싱크와 소스를 설정할 수 있으며 프로퍼티만 수정하여 별도의 스토리지에 데이터를 전송할 수 있도록 한다. Sink에서 실제 수집한 데이터를 특정 위치로 전송하는 과정을 시행하여 이때 단일 Channel과의 커넥션을 연결한다.

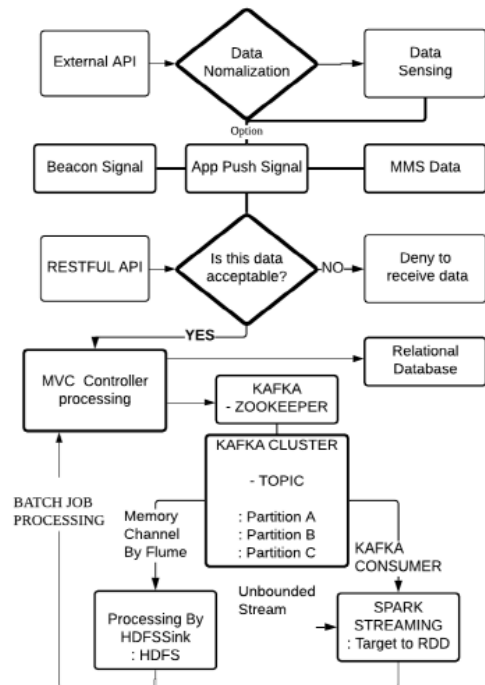


그림 2. 데이터 수집 프로세스의 설계

본 연구에서는 Consolidation Flow를 통해 복합 Sink에서 단일Source로 데이터를 수집하여 제공하는 연동을 제공하도록 하였다. 또한 데이터 스토리지 간에 데이터를 입출력하며 동시에 처리할 수 있는 Spring Cloud Data Flow를 적용하여 효율을 높인다. 본 연구에서는 [그림 2]와 같이 Spring 프레임워크의 Model-

View-Controller (MVC) 컨트롤러에서 내부와 외부 API의 중계형태로 정보를 수집분석하고 KAFKA모듈을 기반으로 Spark Stream와 Flume의 방향으로 데이터를 전달한다. 이후 Spark Stream에서는 실시간으로 전달받은 SNS와 사용자의 특성 정보에 대한 분석 및 처리를 수행하며, Flume모듈에서는 별도의 갱신이 필요한 정보의 배치 처리를 위하여 내부 HDFS에 텍스트 로그를 저장하는 방식으로 전체 시스템을 설계하였다.

2. 시스템의 구현

2.1 시스템의 처리로직

본 연구에서 제안하는 빅데이터 시스템의 로직 구성은 실시간으로 유입되는 이동정보 데이터를 분석하여 해당 고객이 기존에 작성하였던 상점의 평점과 그 상점에 대해 작성하였던 SNS와 상점 방문 정보를 분석과정을 우선 수행 한다. 이 정보를 바탕으로 고객과 유사한 구매 취향을 갖는 고객의 데이터를 대조하여 유사도가 60% 이상인 경우에 해당 고객에서 추천하는 절차를 다시 수행한다. 이때 해당 고객이 이 정보와 다르게 이동정보가 향후에 기록되게 되는 예외상황이 발생하면 이 부분에 대해 피드백 과정이 수행된다. 피드백 과정에서는 실제 고객이 90%이상 방문할 것으로 추측되었으나 실제로는 이동하지 않거나 다른 상점에 방문한 경우를 분석한다. 이 분석을 위해 사용자가 다른 상점에 방문한 이력이 있을 경우, 해당 상점에 정보 요청을 보내서 특정 상황에 대한 것인지에 대한 정보를 수집한다. 추가적인 정보를 수집할 경우, 고객정보 이슈가 발생할 수 있으므로 관계되는 정보는 수집하지 않고 단순한 방문상의 특이사항의 유무만 수집하도록 한다. 이를 통해 사용자들은 사용자입과 동시에 이 시스템을 구성하는 정보 제공자로서 작용하게 되는 것이며 경로상의 정보 제공 알고리즘을 통해 초기에 추천된 정보가 이후에 어떻게 구성될지에 대해 결정하는 후기 과정을 별도로 거치도록 설계하였다.

2.2 상점 정보등록의 처리로직과 추천로직

본 시스템에서 제안하는 상품 정보등록은 [그림 3]과 같이 상점 업주가 직접 정보를 등록한 후 상점과 관계된 정보와 사용자의 피드백 정보가 수집되는 과정을 거

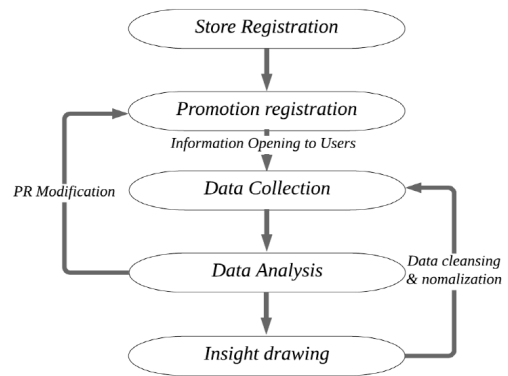


그림 3. 상점 정보의 등록 프로세스

치게 된다. 사용자 데이터 분석과정에서 매출에 도움이 될 수 있는 요소가 확인되면 프로모션을 수정하고 재수행하여 검증과 판매량 증가를 도모한다. 각 상점의 SNS데이터를 분석하고 IBM watson의 감정분석 API에 입력하여 긍정의 비율을 산출하여 이 값을 각 상점에 제공하여 피드백의 한 데이터로서 사용한다. 또한 분석의 결과로 나온 사항은 데이터 정규화에 다시 적용하여 보다 정확한 분석을 유도한다. 예를 들어 “비오는 날에 따뜻한 유자차 할인 프로모션을 저녁에 수행하면 매출 증대에 효과가 있다”라는 결과에 대해 날씨, 기온 인자에 대해 결측치나 오류가 포함된 데이터가 있는지 재확인하고 다른 후보군의 요인이 없는지 확인하고 재분석을 수행한다.

사용자에게 상점 추천에 관련해서는 협업필터링 분석기법을 사용하여 유사한 구매이력을 가진 사용자의 정보를 통해 상점 홍보정보를 푸시, 문자 등의 방법으로 제공하게 된다. 상세 프로세스는 아래와 같다.

- ① 유클리드 거리기법을 통해 사용자를 제외한 전체 사용자의 구매성향에 대한 유사도 테이블을 구성한다.
- ② 상점을 유형화한 테이블을 구성하고 평점 데이터와 SNS 데이터의 수치화한 점수를 포함한다.
- ③ 사용자간의 유사도 분석을 수행하여 추측점수(유사도 * 타인의 평점)를 추정한다.
- ④ 상품 업데이트가 잘 되지 않거나 특정 상품을 재고처리를 위한 특판과 같은 상황에 대해 예외처리를 수

행하여 데이터의 노이즈를 제거한다.

⑤ 추측평점의 총합을 구하여 모든 사람의 평점을 준비한다. 이 데이터를 통해 기존 구매이력이 없는 사람도 사용할 수 있도록 한다.

2.3 연동정보 처리로직

본 시스템에서 제안하는 전체적인 분산처리 구성도와 내외부의 API 연동서버에 대한 설계 방안은 앞서 2장에서 언급한 프로세스 흐름으로 구성하였다.

본 시스템에서는 외부 연동시의 안정성과 편의성을 위하여 Restful API방식, SOAP방식의 두 가지 방식의 제공을 시행한다. 우선 사용자의 스마트폰에서 이동정보와 위치정보를 확인하여 이동할 목적지에 대한 경로정보를 확인한다. 이를 위해 지도와 경로정보에 대한 연동API는 SK telecom의 T map과 NAVER지도를 이용하였다.

실제 사용자의 이동시에 정보를 안정적으로 지원받음과 동시에 중앙처리서버에서 보다 정확한 사용자의 특성 및 경로의 추천상점의 분석이 필요하다. 이를 위해 SNS 와 비콘 정보를 추가적으로 도입하였으며, 특히 SNS정보는 페이스북, 트위터와 같은 일상적인 SNS 정보 뿐아니라, 각 상점에서 제공하는 홈페이지에서의 고객이 남긴 코멘트도 동시에 수집할 수 있도록 각 상점 시스템과의 데이터 연동을 수행하도록 한다. 이를 통해 중앙서버에서는 각 상점의 데이터와 시간단위로 지속적으로 연동하며 이동정보와 동시에 기록된 데이터가 있을 경우에 이를 검증 및 처리 데이터로 활용하게 된다. 또한 사용자로부터 수집할 수 있는 직관적인 데이터에 추가로 웹 크롤링을 통해 수집할 수 있는 각 지역의 다양한 상점과 특성들, 장단점들과 고객의 선호도에 대한 다양한 정보들을 수집하고 정형 데이터 포맷으로 정제하였다.

이 정보들을 일단위로 분석모듈에서 데이터 분석을 수행하여 개별 사용자의 이동데이터에 정보로서 제공할 수 있는 항목에 대해서는 선별하여 향후에 이용할 수 있도록 한다. 이를 위해 정성적인 데이터 분석이 필요하며, 크롤링으로 수집해서 자동 분석된 데이터에 대해 특정 지역의 사용자들에게 리스트에 대한 평점을 요청하고 사용자들에게는 포인트로서 보상을 수행하여

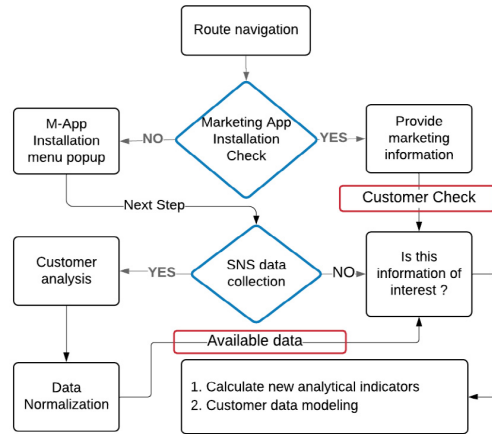


그림 4. 연동API의 내부 프로세스 설계

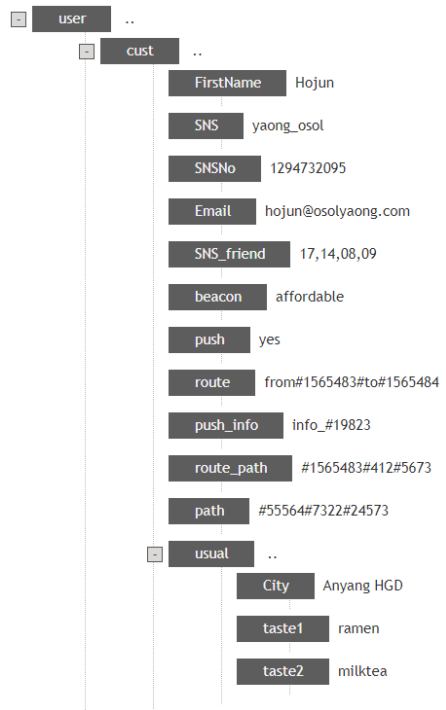


그림 5. XML연동 데이터의 예시

검증을 시행한다. 이를 통해 만들어진 최종적인 상점, 정보리스트를 분석과 제공에 사용한다.

IV. 실효성 검증

1. 실효성 검증의 두 가지 기법

1.1 실효성 검증 기법

본 연구에서 제안하는 시스템의 실제적 효용성을 테스트하기 위하여 통합기능 시뮬레이션 테스트와 심층면접을 시행하였다. 이 두 가지 실험을 수행하여 본 연구에서 제안하는 시스템이 기능적인 측면에서 완성도 높게 작동하고 있으며 이 시스템을 현재의 LBS시스템에 적용가능한지 여부에 대해서도 다각도의 분석이 가능할 것이다. 또한 특징적인 예외처리가 필요한 부분을 확인하고 이를 보완하여 연구적 완성도를 높이기 위하여 이와 같은 검증과정을 수행하였다. 또한 전문가를 대상으로 한 심층면접을 수행하여 실제 기능성 시뮬레이션 테스트에서 확인할 수 없었던 다양한 예외적 이슈들을 사전적으로 검증함과 동시에 연구와 완성도를 높이도록 하였다. 첫 번째로 기능 시뮬레이션 테스트에 있어서 이 시스템은 기본적으로 HTTPS프로토콜 하에서 웹소켓을 기반으로 통신하게 되므로 서버와 어플리케이션간의 데이터 정합성이 중요한 테스트 항목이다. 따라서 통합기능 시뮬레이션 테스트에서는 예외처리 및 오류사항 정비에 초점을 두고 시뮬레이션에 참가한 참가자 5인이 시스템의 사용자로 활동하도록 하였다. 심층면접에서는 시스템 아키텍처 업무를 10년이상 수행한 고급 기술자 등급의 3인을 초빙하여 시스템에 대한 상세한 설명과 이에 대한 분석 및 코멘트를 요청하였고 결과 사항을 시스템에 반영하였다. 3인 중 1인은 지능형 정보시스템 관련 박사학위자이며 나머지 2인은 다수의 대형 정보시스템 프로젝트를 통해 시스템의 안정성과 확장성에 대해 방향성 제시가 가능한 인터뷰어에 해당한다.

1.2 이동성 시뮬레이션 실험

본 연구를 위한 이동성 시뮬레이션 실험에서 효과성을 검증하기 위하여 경기도 안양시 동안구 호계동의 특정지역을 실험경로로 설정하였다. 이 경로는 아파트 단지 2곳이 인접해있고 근처에 빌라지역, 상업지역, 업무지역, 공장지역 등이 개별 블럭으로 구분되어있어서 위치에 따른 지역정보가 다소 정확하게 분석될 수 있다.

또한 개별 지역에 각 상점들이 분산되어 분포되어 있으므로 다양한 예외케이스가 같이 파악될 수 있으며 이동 경로로서도 다수의 지름길을 포함한 경로이므로 본 실험을 위해서 적합한 경로라고 할 수 있다. 또한 이동성 실험을 위해서 해당 경로에 공사, 대로변과 같은 피험자가 위험에 처할 수 있는 경로는 모두 제하였으며 지도API로 사용할 정보의 최신성과 비교하여 실제 거리의 모습과 동일한 것으로 확인하였다.



그림 6. 네이버 지도상의 실험 경로

본 실험에서는 이동성의 정확성에 대해서는 테스트 대상에 포함하지 않으므로 이동경로는 직선으로 단순화 하였으며, 실험 지역에서 테스트할 경로는 빨간색 화살표와 같이 정방향과 역방향에서 각각 실험을 수행하여 가게의 광고판의 방향에 대해서도 입력 정보에 포함되도록 하였다. 또한 가상으로 3곳의 위치를 상점으로 가정하여, 해당 지역에서 SNS정보와 사용자 방문 정보 등을 업로드 하였고 분석처리 데이터로서 가상의 구매이력을 500건을 생성하여 데이터 분석에 활용하였다. 이 준비를 바탕으로 실제 상점 데이터에 준하여 사용자 시뮬레이션을 할 수 있도록 준비하였다. 실제를 가정한 시뮬레이션 테스트로서 테스트 경로에서 날씨와 시간과 같이 광범위한 데이터가 필요로 하는 경우에 대해 정확한 데이터 측정이 어려운 점이 있으므로, 이에 대해서도 날씨와 시간, 근처 인원, 유사성향의 사용자들에 대한 데이터를 수집하기 위하여 실험 참가 인원에 대해 7일간 어플리케이션을 스마트폰에 설치 유지하도록 하였고 이 기간에 걸쳐 정보들을 수집하였다.

따라서 시뮬레이션 테스트가 가질 수 있는 경험적이고 통계적인 예외상황을 보완적인 데이터 수집을 통해 최대한 보충하도록 하였고 실험의 완결성을 높이고자 하였다.



그림 7. 실험 결과의 추천 상점정보

시뮬레이션 실험 수행 이후 결과분석을 시행하였다. [그림 7]과 같이 상점 3곳 중에서 실제 양방향에 따른 경험적인 데이터를 분석하여 파란색으로 표시된 부분이 사용자의 유사한 상점 이용 성향을 갖는 사용자의 SNS등록정보, 주황색으로 표시된 부분이 상점에서의 긍정적인 피드백으로 제공된 것이다. 시뮬레이션 대상자의 기존 이력을 파악하여 성향에 대한 수치를 확인하고, 타 사용자들의 이력 수치와 유사도 검사를 수행하였다. 유사도가 60%이상에 해당하는 사용자 목록을 추출하였다. 이 대상자들을 대상으로 방문가능성이 높은 상점 카테고리에 대한 유사도 행렬을 만들고 각각의 케이스에 대한 예측 점수를 산출한다. 이를 위해 상점 카테고리 5곳(커피, 이불, 미용, 백반, 소형 슈퍼)과 SNS 사용자 중 샘플링한 10인에 대해 데이터 크롤링을 수행하여 30건 이상씩의 실험 데이터를 수집하였다. 실제 상황에서는 특정 상품 카테고리에 대해 이력 데이터가 부족한 경우가 발생할 수 있으며 이렇게 적은 숫자는 추천의 정확도를 하락시킬 가능성이 크다. 그러므로 이런 카테고리의 상황에서는 최소 평가 인원을 정하여 10 이상일때만 유사도를 계산하기로 기준을 설정하여 실험에서 발생가능한 오차를 줄이고자 하였다. 이렇게 아이템과 상점 기반으로 유사도를 계산하여 협업필터

링 추천을 수행하여 사용자에게 정보를 제공하였다. 이 결과로 3곳의 가상의 상점 중 맨 우측의 상점이 선택되었다. 이 결과로 사용자는 우측 맨 제일 오른쪽의 상점에 방문할 확률이 높은 것으로 확인되었으며 상점에 이 정보를 전달하면 상점에서는 보다 적극적으로 홍보를 시행할 수 있게 될 것이다. 또한 실험에서 의도했던 방향성과 관련하여 상점의 입간판을 양방향 중 한 방향에서만 노출되도록 설정하였고 이것을 실제 실험결과에 반영되도록 하였다. 그리하여 나온 결과에서는 실제 주황색과 파란색 화살표가 절반정도로 감소한 결과가 나타났다. 이를 조합하여 상점에서의 마케팅 전략을 세우는 것에 이 데이터를 이용할 수 있음을 알 수 있다.

1.3 심층면접을 통한 분석

시스템 전문가들에 대한 심층면접을 수행하였고 공통적으로 나온 의견은 텍스트 분석에 따른 방법에 대한 것이다. 실제 SNS데이터를 분석하였을 때, 각 고객의 코멘트가 긍정적인지 부정적인지를 정확히 파악하기가 어려우므로 본 연구에서는 특정 감정단어가 코멘트에 포함되어 있는지 여부와 함께 대표적인 동사를 자연어 처리 오픈 API를 활용하여 동사 단위로 분석을 수행하여 가장 유사한 의미의 동사를 산출하는 방법을 사용하였다. 전문가들이 분석한 장점으로는 실제 이동시에 사람마다 움직임의 특징성에 대한 고려를 시스템적으로 구현한 점과 개별상점에서 아주 작은 광고를 시행하더라도 효과를 실시간으로 분석할 수 있는 점에 있다. 반면 시스템 전문가들이 지적한 단점으로는 구매 성향이 유사한 사용자를 분석하는 것에 있어서, 충분한 모수가 확보되지 않으면 정확도가 떨어질 수 있으므로 이에 대한 보다 로직적인 보완과 예외처리가 필요함에 대한 의견이 있었다. 이 실험에서는 예외처리는 시스템의 일반적인 상황과 배치되는 두 가지 상황인 실제 방문으로 예측했지만 반대의 결과가 나온 케이스와 상품이 고객 성향과 상관없는 단순한 특판인 경우가 있다. 이 두 가지 상황의 로직을 확인하기 위하여 상점 카테고리 5곳(커피, 이불, 미용, 백반, 소형 슈퍼)과 SNS사용자의 성향이 정반대인 케이스를 가정하여 시뮬레이션을 수행하였다. 사용자가 방문한 상점에 30분 이후에 푸시 noti를 발송하여 사용자의 구매여부를 입력받도록 하

였고, 특반인 경우에 해당 정보를 기록하도록 하여 예외상황에 대비하는 로직이 정상적으로 구동됨을 검증하였다. 하지만 실제 구매상황에서는 보다 다양한 경우에 예상치 못한 구매가 발생할 수 있으므로 시스템 사용상의 경험이 누적됨에 따라 보다 다양한 예외 케이스의 수립이 필요하다.

위에서 언급한 내용과 같이 본 시스템에서 보이는 장점과 개선이 필요한 사항들이 도출되었다. 긍정적으로 보고된 항목은 본 시스템 설계의 목표와 합치되는 것으로 확인되었으며 개선사항으로 도출된 항목은 추가적인 정보 평가 및 추천 모듈의 개발을 통해 보완이 가능한 것으로 확인하였다.

V. 결론

본 연구에서는 빅데이터 시스템을 활용하여 각 지역 상점들이 마케팅에 활용할 수 있는 정보를 창출함과 동시에 사용자들이 본인들의 요구에 맞는 상점을 쉽고 빠르게 추천받을 수 있는 융합적인 정보시스템의 설계를 목적으로 하였다. 본 연구에서는 무엇보다 사용자와 마케팅의 두 가지 관점에서 모두 만족할 수 있도록 목적 달성의 측면과 시스템 안정성의 양 측면을 추구하고자 하였다. 이를 위해 빅데이터 분산처리 시스템을 도입하였으며 전체 정보의 처리 속도와 시스템상의 안정성을 높일 수 있었다. 또한 사용자의 이동정보를 로그로서 수집하는 본 시스템의 특성상 사용자가 증가하면 데이터양이 폭발적으로 증가하게 되므로 본 시스템의 상용 서비스로의 활용을 위해서는 데이터의 분리 보관정책 혹은 데이터 정제 기법의 추가 보완이 반드시 필요하다. 본 연구의 한계점으로 사용자가 특정 경로를 이동한다는 전제하에 방문가능성이 있는 상점을 대상으로 하였으므로, 특정 상점을 방문하기 위해 특정 경로를 이동한 경우에 대한 예외적인 데이터 보완이 후속 연구에서 필요하다. 이와 같은 연구의 보완을 통해 상점의 관점에서는 보다 정확한 판촉 전략을 수립할 수 있을 것이고 고객 입장에서는 보다 편리하게 본인에게 잘 맞는 상점을 찾을 수 있도록 다양한 정보를 손쉽게 제공할 수 있게 되므로 데이터의 축적에 따라 점차 진화

적 어플리케이션으로 발전되어갈 수 있을 것이다.

참고 문헌

- [1] <https://www.technavio.com/>, 2020.1.12.
- [2] 한국인터넷진흥원, *2017 LBS 산업 실태조사 보고서*, 2017.
- [3] 한국인터넷진흥원 보고서, *국내외 LBS 산업 동향 보고서*, 2015.5.28.
- [4] <https://www.boannews.com/>, 2020.1.12.
- [5] <https://www.hankyung.com/>, 2020.1.12.
- [6] <https://www.daangn.com/>, 2019.12.30.
- [7] 이윤주, 서지훈, 최진탁, “국내 관광 서비스 제고를 위한 안드로이드 기반의 위치정보 통합 시스템 구축 기법에 관한 연구,” *융복합지식학회논문지*, 제1권, 제2호, pp.71-78, 2013.
- [8] 한국마케팅연구원 편집부, “위치 기반 마케팅 (Location Based Marketing),” *마케팅지*, 제49권, 제12호, pp.60-66, 2015.
- [9] 손동진, 김혜경, “2017소셜 인플루언서를 활용한 디지털 마케팅 전략 연구: 올레드TV 글로벌 디지털 캠페인 사례를 중심으로,” *광고 PR실학연구*, 제10권, 제2호, pp.64-95, 2017.
- [10] C. B. Mun and O. Lee, “VR-based Location Aware System Modeling: Providing Location Information for Urgent Requests,” *International Journal of Applied Engineering Research*, Vol.12, No.15, pp.4955-4958, 2017.
- [11] 문창배, *Integrated Supporting Platform for the Visually Impaired: Using Smart devices and Web-based Server*, 한양대학교, 박사학위논문, 2018.
- [12] 문창배, 박현석, “빅데이터 분산처리 시스템을 활용한 지능형 LBS서비스의 설계,” *한국콘텐츠학회논문지*, 제19권, 제2호, pp.159-169, 2019.
- [13] 김은택, 김정빈, 금경조, “인스타그램 위치정보 데이터를 이용한 읍지로 3·4가 지역 활성화의 실증분석,” *서울도시연구*, 제20권, 제2호, pp.19-35, 2019.
- [14] 김태현, 이재길, 윤수식, “스마트카드 빅데이터 분석을 통한 사람들의 미래 위치 예측,” *한국정보과학회 한국소프트웨어융합학술대회 논문집*, pp.948-950, 2017.

저 자 소 개

문 창 배(Chang-Bae Mun)

정회원



- 2011년 2월 : 한양대학교 정보통신 학부 컴퓨터 전공(공학사)
- 2018년 2월 : 한양대학교 정보시스템학과(공학박사)
- 2019년 11월 ~ 현재 : 한양사이버대학교 전기전자통신공학부 조교수

〈관심분야〉 : 지능형 플랫폼, 기술경영

박 현 석(Hyun-Seok Park)

정회원



- 2008년 2월 : 한양대학교 미디어통신공학과(공학사)
- 2014년 8월 : 포항공과대학교 기술경영대학원(공학박사)
- 2016년 3월 ~ 현재 : 한양대학교 정보시스템학과 조교수

〈관심분야〉 : 기술혁신 방법론, 기술 경영