



빅 데이터기반 마이닝 마인즈 헬스케어 프레임워크

저자 (Authors)	Oresti Banos, Wajahat Ali Khan, Muhammad Bilal Amin, 허태호, 방재훈, 강동욱, Maqbool Hussain, Muhammad Afzal, Taqdir Ali, 이승룡
출처 (Source)	한국통신학회지(정보와통신) 32(11) , 2015.10, 12-20 (9 pages) The Journal of The Korean Institute of Communication Sciences 32(11) , 2015.10, 12-20 (9 pages)
발행처 (Publisher)	한국통신학회 Korea Institute Of Communication Sciences
URL	http://www.dbpia.co.kr/Article/NODE06531551
APA Style	Oresti Banos, Wajahat Ali Khan, Muhammad Bilal Amin, 허태호, 방재훈, 강동욱, Maqbool Hussain, Muhammad Afzal, Taqdir Ali, 이승룡 (2015). 빅 데이터기반 마이닝 마인즈 헬스케어 프레임워크. 한국통신학회지(정보와통신), 32(11), 12-20.
이용정보 (Accessed)	경희대학교 163.180.116.67 2015/12/08 11:17 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다.

이 자료를 원저작자와의 협의 없이 무단게재 할 경우, 저작권법 및 관련법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

The copyright of all works provided by DBpia belongs to the original author(s). Nurimedia is not responsible for contents of each work. Nor does it guarantee the contents.

You might take civil and criminal liabilities according to copyright and other relevant laws if you publish the contents without consultation with the original author(s).

빅 데이터기반 마이닝 마인즈 헬스케어 프레임워크

Oresti Banos, Wajahat Ali Khan, Muhammad Bilal Amin, 허태호, 방재훈, 강동욱,
Maqbool Hussain, Muhammad Afzal, Taqdir Ali, 이승룡

경희대학교

요약

최근 의학 기술이 눈부시게 발전함에 따라 사람들은 수명이 연장되고 삶의 질 향상에 많은 관심을 가지게 되었다. 더욱이 혁신적인 디지털 기술 발전과 함께 다양한 웨어러블 기기와 수많은 헬스케어 어플리케이션이 출시되고 있으며, 이들은 어떻게 하면 개인의 성향이나 체질에 잘 맞는 맞춤형(개인화) 서비스를 제공할 수 있을 것인가에 관심을 두고 진화하고 있다. 따라서 IoT 환경의 일상생활에서 입력되는 센서 데이터의 수집, 처리, 가공 기술, 일상 행위 및 라이프 스타일 인지, 지식 획득 및 관리 기술, 개인화 추천서비스 제공, 프라이버시 및 보안을 통합적으로 지원할 수 있는 프레임워크 개발에 대한 요구가 증대되고 있다. 이에 본 고에서는 저자가 개발중인 개인 맞춤 건강 및 웰니스 서비스를 제공하는 마이닝 마인즈 프레임워크를 소개한다. 마이닝 마인즈는 현존하는 최신 기술의 집약체로 개인화, 큐레이션, 빅 데이터 처리, 클라우드 컴퓨팅의 활용, 다양한 센서 정보의 수집과 분석, 진화형 지식의 생성과 관리, UI/UX를 통한 습관화 유도 등 다양한 요소를 포함한다. 그리고 건강 및 웰니스 프레임워크 요구사항 분석을 통해 마이닝 마인즈가 이러한 요구를 충족시킬 수 있으며, 개발된 프로토타입을 통해 개인화 서비스의 발전 가능성을 입증하고 향후 나아가야 할 방향을 제시한다.

키워드 : 디지털 건강, 개인화, 큐레이션, 빅 데이터, 클라우드 컴퓨팅, 행위 인지, 상황 인지, 지식 베이스, UI/UX, 프라이버시, 보안

I. 서론

최근 건강 관리 서비스의 주요 트렌드는 질병 발병 후의 사후 관리가 아닌 질병의 사전 예방과 건강 관리를 목표로 언제, 어디서, 어떻게 특정 환자 및 사용자에게 맞춤형 서비스를 제공해야 할 것인지에 초점이 맞추어져 있다[1][2]. 일반적으로 알려진 만성질환은 해로운 식 습관 및 생활 습관, 음주, 흡연, 운동부

족, 스트레스 등 잘못된 라이프 스타일의 결과에 주로 기인된다 [3][4][5]. 따라서, 개인의 건강한 생활습관 유도과 질병 및 장애에 대한 대처를 보다 효과적으로 이끌어낼 수 있는 혁신적인 헬스케어 기술 개발이 요구된다.

정보통신 기술은 새로운 혁신적 헬스케어 시대의 도래를 가능케 하는 핵심으로, 사람들이 개인의 건강 상태를 제대로 파악하여 적시적소에 건강 정보 및 서비스를 지원받을 수 있도록 한다 [6]. 최근 모바일 및 웨어러블 기술의 발전으로 인하여 건강 관리 어플리케이션과 시스템 개발에 대한 관심이 고조되고 있다 [7]. 이러한 기술의 주된 목표는 개인의 건강한 라이프스타일 유지를 위한 서비스 제공으로, 특히 신체 건강과 관련되어 사용자의 일상생활을 추적하며 건강 유지를 위한 동기를 부여한다. 예를 들어, Withings의 Activite, Garmin의 Vivofit 및 Misfit의 Misfit Shine 등은 최근 팔찌나 시계 형태의 웨어러블 기기를 개발, 상용화하여 판매하고 있으며, 이를 모바일 앱과 연동시켜 사용자의 걸음걸이 수 및 수면 시간을 측정하여 건강 유지를 위한 기본적인 추천 서비스를 제공한다. 이보다 한 단계 더 발전된 서비스로는 신체 건강 상태에 대한 경고[8]와 만성질환 감지[9] 등이 있다. 이러한 시스템 중 일부는 건강 관리를 위한 교육적인 요소와 개인 코치 기능을 제공하여 건강 생활을 유지할 수 있도록 한다[10]. 이렇듯 건강 관리 시스템은 진일보하였으나 시스템 성능의 한계, 서비스 제공 범위의 제약, 상호운용성의 결여 등, 한계점이 존재한다.

한편, 어플리케이션에 종속적이지 않고 자원 관리의 효율성 향상 및 건강 정보 시스템을 널리 활용하기 위한 프레임워크 개발이 이루어지고 있다. [11]에서는 다양한 인터페이스와 다중 파라미터 모니터링을 통합한 미들웨어 프레임워크를 제안하였으며, [12]에서는 센서 데이터의 분석 및 분류를 분산 신호 처리 알고리즘을 이용하여 빠른 처리가 가능한 신체 센서 네트워크 프레임워크를 제안하였다. [13]에서는 사용자의 행위, 심리, 생체 정보를 수집하는 모바일 플랫폼을 제안하였고, [14]에서는 사용자의 몸에 부착된 생체 신호와 가전기로부터 얻은 데이터 기반으로 사용자의 일일 약물 관리 헬스케어 플랫폼을 제안하였다. [15]에서는 자원 및 통신의 간결화, 웨어러블 건강 데

이터의 수집 및 지식 추출을 지원하는 독자적인 프레임워크를 제안하였고, [16]에서는 시스템 개발자들이 디지털 건강 데이터를 유용하게 쓸 수 있도록 도와주는 오픈 모바일 프로젝트를 제안하였다. Google의 Google Fit, 삼성의 SAMI 및 Apple의 HealthKit은 모두 사용자의 건강 데이터를 수집하여 다양한 건강 및 웰니스 어플리케이션 간에 데이터의 통합과 정보 공유를 지원하는 대표적인 상업 플랫폼이다.

이렇게 개발된 플랫폼들은 건강 및 웰니스 서비스 발전에 대한 공헌을 하였으나, 여전히 개선의 여지가 있다. 예를 들어, 대부분의 모바일 건강 프레임워크는 스마트폰 기반으로 사용자 어플리케이션의 지속적인 관리와 업데이트가 필수적이며, 수집된 사용자 데이터는 해당 단말기에서만 보관되어 다른 시스템에서의 사용이 불가능하다. 또한, 다양한 시스템 및 어플리케이션의 사용은 비슷한 종류의 데이터 수집과 결과를 도출하여 불필요한 중복성과 과도한 연산을 야기시킨다. 이러한 시스템들은 사용자로부터 요청이 발생할 때만 동작하여 중요한 건강 관련 정보를 놓칠 가능성도 높으며, 클라우드를 이용한 데이터 공유와 통합을 지원하도록 고안된 시스템들은 대부분 단순히 저장 용도로만 사용되고 있다.

이러한 제약점을 보완하기 위해 본 고에서는 개인화된 건강 및 웰빙 서비스 제공을 통하여 건강한 라이프스타일을 지속할 수 있도록 도와주는 분산 마이닝 마인즈 프레임워크 개발 경험을 소개한다. 이 프레임워크는 다양한 멀티모달 입력 소스로부터 얻은 일상 생활 데이터를 실시간으로 마이닝하여 사용자의 행위 및 라이프스타일 인지하고, 헬스케어 지식획득 및 지식베이스의 지속적인 관리와 추천서비스, 그리고 UI/UX 통하여 사용자에게 건강에 대한 관심과 생활 습관변화를 유도할 목표로 한다. 또한 클라우드를 단순히 저장 목적만으로 사용하는 것이 아닌 고성능의 컴퓨팅 능력을 활용하여 사람들의 건강 및 동향을 파악하고, 이를 토대로 각 개인에게 맞춤형 건강 관리 서비스를 제공한다.

II. 개인 맞춤형 헬스케어

1. 프레임워크의 요구사항

미래의 헬스케어 프레임워크에 빅 데이터 처리 기술 도입은 필수 불가결하며, 빅 데이터의 요구사항인 4V (Variety(다양성), Volume(양), Velocity(속도), Veracity(진실성))를 충족할 수 있어야 한다. 기술의 발전에 따라 현재 웨어러블 기기, 소셜 네트워크, 의료 시스템 등을 통해 다양한 데이터의 수집이 가

능해졌으며, 가공된 정보를 이용하면 사람의 건강 상태를 표현하는 것이 가능하다. 이러한 정보에는 센서 데이터를 통한 신체 정보, 사용자 프로파일 및 관심도를 통한 논리 정보, 소셜네트워크를 통한 사회성 정보, 의료 데이터를 통한 임상 정보 등이 있다. 따라서, 다양한 빅 데이터와 정보를 지속적으로 관리할 수 있는 새로운 형태의 디지털 헬스 프레임워크가 필요하며, 기존의 수기로 기록된 진료 기록이나 멀티미디어 등의 비정형 데이터를 통합하여 시스템에서 활용이 가능하도록 정형 데이터로 바꿀 수 있는 기능이 필요하다. 새로운 센싱 및 멀티미디어 기술 개발로 인하여 단일 사용자가 하루 동안 생성하는 데이터의 양은 규모가 상당하며, 여러 사용자로부터 매일 데이터를 수집할 경우 그 규모는 기하급수적으로 늘어나게 된다. 이렇게 방대한 데이터의 보관 및 사용의 용이성을 위해서는 새로운 연산 및 저장 기술이 요구된다. 또한 다양하고 많은 데이터가 지속적으로 실시간으로 유입되기 때문에 빠른 처리 속도도 중요한 요구사항이다. 마지막으로 서로 다른 종류의 데이터 형식으로 인한 데이터 불필요성, 중복, 충돌을 해결하는 것도 중요한 요소이다. 이에 따라, 데이터의 진실성을 위해 일관성, 의미성, 투명성 및 정확성을 다룰 수 있어야 한다.

획득된 데이터 중 건강과 관련된 의미 있는 정보를 추출하는 것은 단순 수집 작업보다 훨씬 더 고난도의 기술이 요구된다. 따라서, 디지털 헬스케어 프레임워크는 이를 자동으로 처리할 수 있는 지능적인 메커니즘을 보유해야 한다. 프레임워크는 인간 중심의 데이터 처리, 해석 가능한 정보와 통찰력을 추출할 수 있어야 하며, 이를 토대로 개인 맞춤형 건강 서비스를 지원할 수 있어야 한다.

건강 및 웰니스 정보와 지식은 고품질의 헬스케어 서비스 제공을 위해 사용되어야 한다. 시스템은 사용자에게 이해하기 쉬운 형태의 서비스를 알람, 추천 및 가이드라인 형식으로 제공해야 한다. 그러나, 현재 대부분의 디지털 헬스케어 시스템은 사용자에게 동일한 서비스를 반복적으로 제공하여 각각의 개인이 가진 특성을 반영하지 못한다. 따라서, 미래 헬스케어 프레임워크의 중요한 요구사항 중 하나로 시스템이 인간 중심의 개인화 서비스를 제공할 수 있어야 한다. 이를 위해서는 전문가 시스템이 필요하며, 이는 각 사용자에게 맞는 최적의 추천 정보 제공과 그에 대한 설명 정보를 쉽게 제공할 수 있도록 한다.

헬스케어 시스템을 이용하는 사용자들은 다양한 상황에 놓여 있을 수 있다. 예를 들어 한시가 급한 위중 환자의 경우 그들의 건강 상태를 한 눈에 빠르게 볼 수 있길 원하며, 개인의 건강 상태를 중요시하는 사용자는 그들의 건강 상태를 세부적으로 알길 원하고, 임상 전문가들은 다양한 사람들의 건강에 대한 상태를 상세히 알기 원한다. 따라서, 사용자 인터페이스는 다양한

사용자들에 알맞은 맞춤형으로 구성될 필요가 있다. 사용자의 경험 또한 중요한 고려 요소이므로, 인터페이스는 사용자의 개인 경험을 고려하여 유용성, 편리성, 효율성을 제공해야 한다. 사용자의 경험은 시간과 상황에 따라 쉽게 변화하기 때문에, 사용자의 응답과 행위를 지속적으로 관찰하여 적지적소에 맞게 인터페이스를 변화하여 제공해야 한다.

서비스 제공에 있어 보안 및 프라이버시는 반드시 이루어져야 한다. 일반적으로 사용자들은 자신들의 개인 정보가 공개될 원하지 않으며, 이에 따라 보안은 민감한 개인정보 보호를 위해 필수적이다. 데이터와 서비스가 분산 환경에서 제공되는 경우에는 더 강도 높은 보안 기술이 요구된다. 데이터의 소유권, 악의적인 사용, 규제, 법 정책 등은 이러한 보안 요소를 포함한 헬스케어 시스템의 사용에 중요한 장애 요소이므로, 디지털 헬스케어 프레임워크는 프라이버시, 보안, 보호, 리스크 관리가 무엇보다 중요하다.

Ⅲ. 마이닝 마인즈 프레임워크

본 고에서는 미래의 디지털 헬스케어 프레임워크의 요구사항을 반영하는 독자적 프레임워크인 마이닝 마인즈를 제안한다 (그림 1).

마이닝 마인즈란 다양한 종류의 데이터를 수집하여 사용자의 일상 생활 정보를 추론하고, 이를 기반으로 개인화된 웰빙 및

헬스케어 서비스를 제공하는 혁신적인 서비스, 툴 및 기술의 집합이다. 마이닝 마인즈의 기본 철학인 큐레이션은 데이터, 정보, 지식, 서비스로 나뉘며, 빅 데이터 분석, 지식의 획득과 유지보수, UI/UX를 통한 개인의 습관화 유도 기능을 내포한다. 데이터 큐레이션은 다양한 멀티모달 센서로부터 얻은 데이터의 처리와 저장을 담당하며, 데이터 종류로 스마트 폰 및 웨어러블 기기를 통한 가속도, GPS, 소셜 네트워크, 설문, 영상, 오디오, EMR/EHR/PHR 과 같은 의료데이터 등이 포함된다. 데이터 큐레이션에서 처리된 데이터는 정보 큐레이션에서 사용되어 저수준 및 고수준 컨텍스트 인지를 통해 사용자의 행위, 상황과 함께 행위패턴 및 라이프스타일을 구성하는 신체적, 정신적, 사회적 영향을 추론한다. 이들은 지식 큐레이션에서 웰빙 및 건강 지식을 생성하고 진화시키기 위해 사용된다. 각 계층의 데이터, 정보, 지식은 서비스 큐레이션으로 이동되어 개인 맞춤형 건강 추천 서비스를 생성한다. 지원 계층의 보안 및 프라이버시를 통해 모든 데이터 처리 및 서비스에 대해 개인 정보 보호를 강화하며, 사용자 경험, 피드백, 동향 등을 분석한다.

1. 데이터 큐레이션 계층 (Data Curation Layer: DCL)

DCL은 다양한 형태의 소스로부터 데이터를 획득, 큐레이션 및 저장하는 역할을 담당하며 크게 빅 데이터 처리와 빅 데이터 저장소 컴포넌트로 나뉜다. 빅 데이터 처리 모듈에서는 데이터의 획득, 레이블링과 분석을 담당하며, 상황에 따라 실시간 혹은 비실시간 방식으로 이루어진다. 다양한 종류와 방대한 양의 데이터가 실시간으로 입력되기 때문에 데이터의 노이즈 및 중복성이 많다. 따라서 DCL에서는 빅 데이터 저장소에 데이터를 저장하기 전 데이터 레이블링 과정을 실행 시간에 실행하고 다양한 종류의 분석 필터를 이용하여 데이터의 일관성 및 신뢰성을 유지하고 이 후 데이터를 성격에 맞게 분류하고 분석한다. 실시간 데이터 필터링 이외에도 기존에 저장된 데이터의 특징에 따라 오프라인 일괄처리 방식의 필터링도 진행된다.

저장된 데이터들은 추후 활용을 위하여 삭제되지 않는다. 이 과정에서 지원 계층(Supporting Layer: SL)의 보안 및 프라이버시 컴포넌트가 사용자 인증과 암호화 역할을 수행한다. 데이터의 보다 쉬운 활용을 위하여 라이프로그 온톨로지를 이용하여 지식의 추출과 데이터 기반 서비스 프로비저닝을 통해 데이터를 고수준으로 표현한다. 이러한 정보는 인터미디어트(Intermediate) 데이터베이스에 저장되어 서로 다른 계층과 어플리케이션 간에 쉽고 빠른 데이터 공유가 가능하도록 한다. 이 과정에서 수집된 비정형 스트림 데이터는 반정형 데이터 형태로 변형되어 라이프로그 온톨로지 매핑된다.

ICL(information Curation Layer)에서 생성한 저수준 및

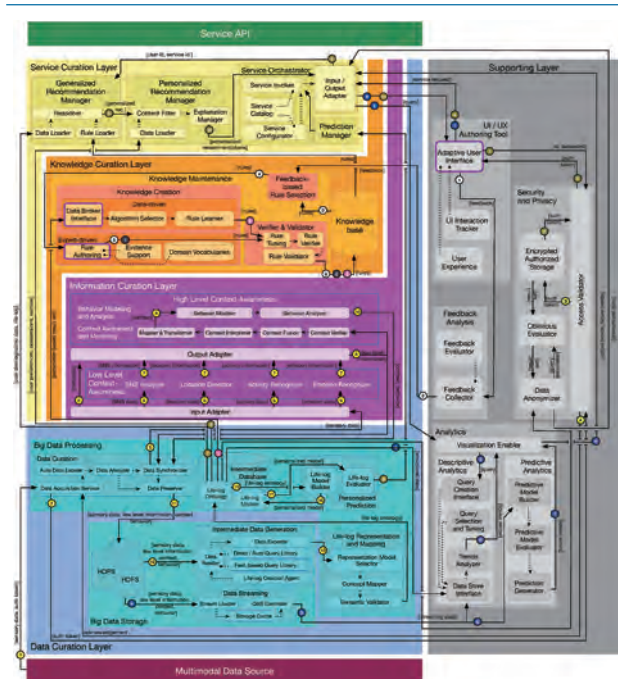


그림 1. 마이닝 마인즈 프레임워크 아키텍처

고수준 상황인지 모델과 DCL의 라이프로그 온톨로지를 이용하여 사용자의 일상 생활 데이터와 정보를 시간순서대로 정형 데이터 형태로 저장하며, 향후에 KCL(Knowledge Curation Layer)과 SCL(Service Curation Layer)에서 사용된다. DCL은 세부 모듈인 개인화 예측(Personalized Prediction)에서 각 사용자의 행위에 따른 사전 예측을 진행하며, 이는 라이프로그 모델로 표현되어 SCL에서 최종 예측을 위한 정보로 활용되며 개인화된 추천 정보를 생성한다. SL의 시각화와 분석 서비스를 위해 데이터 스트리밍 컴포넌트가 스트림 데이터에 대한 동기화를 담당한다.

2. 정보 큐레이션 계층 (Information Curation Layer: ICL)

ICL은 크게 고수준 상황인지(High Level Context Awareness HLCA)와 저수준 상황인지(Low Level Context Awareness, LLCA)로 구성되며, 사용자의 행위와 상황을 추론하고 모델링하는 역할을 담당한다.

LLCA는 사용자가 실생활에서 발생시키는 다양한 종류의 데이터인 신체 활동, 감정 상태, 위치, 소셜 네트워크 사용 패턴 등을 인지하여 카테고리화 하며, 이러한 데이터는 HLCA에서 융합되고 추론을 통해 최종 사용자의 행위 및 상황을 추론한다. LLCA는 SNS 분석기, 행위 인지, 감정 인지 및 위치 탐지기 컴포넌트로 이루어진다. SNS 분석기는 사용자가 페이스북이나 트위터와 같은 소셜 네트워크 서비스를 사용하면서 열람한 목록, 댓글, 방문 기록 및 동향에 대한 텍스트와 멀티미디어 데이터를 수집하여 분석한다. 이를 통해 개인의 관심 영역, 요구사항, 상태 등을 판별할 수 있다. 사용자의 전반적인 신체 활동은 행위 인지모듈을 통해 인지된다. 사용자의 행위는 몸에 부착된 관성 센서, 카메라 영상, 오디오 등의 데이터를 이용하며, 행위 인지 결과로 서기, 걷기, 뛰기 등의 행위가 인지된다. 감정 인지 모듈도 같은 방법으로 다양한 종류의 데이터를 수집하여 기쁨, 슬픔 등의 감정 결론을 도출한다. 사용자의 상황을 정확히 파악하기 위해서는 사용자의 현재 위치나 이동 경로를 파악하는 것이 중요하며, 이는 위치 탐지기 모듈에서 담당하며 GPS 정보를 이용한다. LLCA의 4가지 컴포넌트에서 요구되는 데이터의 종류 선택은 입력 어댑터(Input Adapter) 컴포넌트가 담당하며, DCL로부터 큐레이션 된 데이터를 로드 한다. LLCA에서 생성된 정보는 출력 어댑터(Output Adapter)를 통해 DCL에 전송된다.

LLCA에서 분류된 데이터는 HLCA에서 사용자에게 대한 보다 정확하고 상세한 상황인지 추론에 사용된다. HLCA의 컴포넌트는 두 가지로 구성되며, 첫번째는 상황 인지 및 모델링(Context Awareness & Modeling) 컴포넌트로 사용자의 상황

을 표현하고 해석한다. 일반적으로 널리 사용되는 모델링 방법인 온톨로지를 이용하여 사용자의 상황을 표현하고 추론하며, 추론은 룰 기반으로 진행된다. 예를 들어 사용자가 특정 시간에 식당에서 앉아 있을 경우 현재 식사를 하는 것으로 추정할 수 있다. 상황인지 모델링의 핵심요소는 라이프로그 저장소로, 여기에 사용자의 모든 상황 정보가 저장된다. 이는 HLCA의 두번째 컴포넌트인 행위 모델링 및 분석(Behavior Modeling and Analysis)에서 사용되며, 사용자 행위 패턴과 동선을 파악하는 역할을 담당한다. 예를 들어 특정 사용자가 특정 시간대에 점심을 먹으러 가는 경우 이는 해당 사용자만의 개인화된 행위 패턴 혹은 동선이라 할 수 있다.

3. 지식 큐레이션 계층 (Knowledge Curation Layer, KCL)

KCL의 핵심 역할은 건강 및 웰니스 지식의 생성과 관리이다. 지식은 데이터 기반 혹은 전문가 기반 접근법을 이용하여 도메인 전문가나 지식 엔지니어에 의해 생성된다. 전문가 기반 접근법의 경우 룰 저작 컴포넌트를 이용하여 전문가가 손쉽게 룰을 생성할 수 있도록 하며, 생성된 룰의 신뢰성을 위하여 기반 증거 자료 수집과 문법 교정기를 이용한 오류 제거 기능을 보유한다. 데이터 기반 접근법은 DCL과 ICL로부터 큐레이션 된 라이프로그 정보를 이용하며, 데이터 브로커(Data Broker) 인터페이스에서는 전문가가 선정한 기준이나 요인에 따라 인터미디어트(Intermediate) 데이터베이스로부터 필요한 정보만을 수집하며, 이 과정은 다양한 마이닝 모델을 이용하여 설정된 요인과 관계 있는 정보만을 자동으로 수집한다. 생성된 룰은 일관성을 위한 검증과 룰 형식 위반, 중복 등을 방지하기 위한 인증 과정을 거친다.

지식의 진화는 두 가지 핵심 과정을 통해 이루어진다. 첫번째는 지식의 생성 과정으로 전문가가 입력한 새로운 건강 및 웰니스 지식과 새로운 라이프로그 내용에 따라 룰은 동적으로 업데이트되거나 교체된다. 두번째 과정은 피드백 정보의 활용으로, 이는 기존의 지식을 검증하는 과정에서 중요하게 사용된다. 예를 들어 피드백은 SCL이 제공한 추천 정보에 대한 사용자들의 만족도를 반영하는데, KCL은 SL에서 처리되어 제공된 피드백 정보를 이용하여 룰 간의 충돌을 감지하고, 전문가가 룰 저작 컴포넌트를 통해 룰의 추가적인 분석, 수정 및 검증을 할 수 있는 바탕 자료로 활용한다.

4. 서비스 큐레이션 계층 (Service Curation Layer: SCL)

SCL은 DCL, ICL, KCL에서 큐레이션 된 데이터, 정보, 지식을 사용자에게 제공되는 실제적인 건강 및 웰니스 지원 서비스로 변환한다. 서비스 오케스트레이터(Service Orchestrator)

가 전체적인 서비스 관리를 담당하며, 서비스 요청에 대한 응답, 필요한 서비스의 로딩 및 서비스 큐레이션에 필요한 서비스의 프로세스 조정 등의 역할을 수행한다. 서비스 요청의 종류에는 정시에 의한 요청, 이벤트에 의한 요청과 사용자에게 의한 직접적인 요청이 있으며, 이를 만족시키기 위해 필요한 서비스를 호출한다. 제공되는 주요 레퍼런스 서비스는 개인 맞춤형 건강 및 웰니스 추천 정보의 생성이며 두 가지 단계로 구성된다. 첫 단계는 DCL이 제공한 데이터와 KCL의 지식을 토대로 사용자 통계 정보와 라이프로그를 이용한 추론을 통하여 일반적인 추천 정보를 생성한다. 두번째 단계는 생성된 일반적인 추천 정보를 개인에 맞게 맞춤화 시키는 단계로 사용자의 선호도, 신체 상태, 주변 환경 및 요구사항 정보를 토대로 추천 정보를 가공한다. 예를 들어 어느 사용자의 건강 증진을 위해 운동을 장려하는 경우, 사용자가 자전거를 소유하지 않으면 자전거 운동을 추천하지 않으며, 사용자가 출장 중일 경우 평소에 가던 운동 센터에 갈 것을 추천하지 않는다. 개인화된 추천 정보는 사용자가 이해하기 쉽도록 정제되며, 사용자의 관심을 유도하기 위한 동기부여 정보를 포함시킨다.

추천 정보 생성 과정은 예측 매니저 (Prediction Manager)를 통해 시작된다. 이 컴포넌트는 사용자의 라이프 패턴을 분석하여 미래에 발생 가능한 행위를 예측하며, 이 때 필요한 서비스의 실행 및 추천 정보를 생성한다.

5. 지원 계층 (Supporting Layer: SL)

SL계층은 데이터에 대한 심화 분석, 사용자와의 상호작용, 개인화된 UI/UX, 목시적 및 명시적 피드백 분석, 프라이버시 및 보안 요소를 지원하여 마이닝 마인즈 프레임워크를 전체적으로 보조하는 역할을 한다.

분석 모듈은 수집된 데이터 및 큐레이션 데이터를 통해 사용자들의 건강 상태와 행위 패턴 및 동향을 분석한다. 이러한 동향 정보에는 현재의 사실이나 미래의 동향 내용이 포함되며, 통계 분석 그래프를 통해 보여준다. 예를 들어 일상 생활 데이터를 추론하여 청소년들의 수면과 스트레스 사이의 패턴을 찾아낼 수 있다. 미래의 예측 사실은 예측 분석(Predictive Analytic)에서 담당하며, 이는 기술적 분석 (Descriptive Analytic)에서의 결과를 이용하여 회귀 및 기계학습 기법을 사용하여 예측한다. 분석된 내용은 시각화를 통해 보여주며, 사용자의 직업이나 성향에 따라 시각화 내용을 달리한다.

사용자의 선호도, 취미, 상태에 따라 적응형 UI/UX 모듈은 스마트폰의 최종 사용자 어플의 형태를 지속적으로 변환하며, 폰트 크기, 테마 배경, 음성 크기 등의 변환이 포함된다. UI/UX는 또한 피드백 정보를 수집하는데 중요한 역할을 한다. 사

용자가 단말 화면을 통해 입력한 명시적 정보뿐만 아니라 사용자의 행동 반응을 탐지하는 암시적 정보를 토대로 사용자에게 제공된 추천 정보의 만족도를 알 수 있으며, 이러한 정보는 지식의 진화와 유지보수에 사용된다.

사용자로부터 수집된 데이터는 민감한 개인 정보를 포함하므로, 보안과 프라이버시 컴포넌트는 사용자 데이터 수집과 처리의 전 과정에 걸쳐 데이터의 암호화와 보안 프로토콜을 사용한다. 저장소 보안을 위해서 AES 표준이 사용되며, oblivious processing을 위해 homomorphic 암호화와 private matching 기법이 사용된다. 마이닝 마인즈 프레임워크와 최종 사용자 어플리케이션 간의 빈번한 데이터 교류로 인하여 확률화 기법을 사용하여 데이터의 최소 누설과 고 엔트로피를 보장하며, 인증 모델 사용으로 인증된 사용자에게만 정보를 공개하고, 암호화로 인한 데이터 처리 속도 저하를 해결하기 위해 민감 정보와 민감하지 않은 정보를 구분한다.

IV. 마이닝 마인즈 프레임워크 구현

본 절에서는 마이닝 마인즈 프레임워크의 초기 구현 버전에 대해 소개한다. 마이닝 마인즈는 체중 관리를 통한 개인의 건강한 라이프스타일 유지를 유도한다. 마이닝 마인즈는 클라우드 환경에서 동작하는 분산 플랫폼으로 데이터에 대한 보존과 강력한 연산 능력을 보유하며, 마이크로소프트의 Azure 퍼블릭 클라우드에 배치되었으며 각각의 계층은 독립된 표준 가상 인스턴스로 존재하며 게스트 OS로 윈도우 서버 2012 R2을 사용한다. SL은 독립적인 인스턴스로 존재하지 않고 각 계층의 일부 형태로 배치된다. 계층별 배치는 캡슐화로 인한 보안성 증가와 재사용이 가능하다는 장점을 가지며, 계층 간의 통신은 DCL의 웹 서비스가 관장하며, SQL 서버를 이용하여 인터미디어트 데이터베이스를 주축으로 요청된 정보에 따라 DCL에 의한 데이터 교환이 이루어진다. 교환 데이터 중에는 데이터 모델이 포함되며, 객체 모델로써 각 계층간에 공유된다.

동적 라이프 스타일 서비스 지원을 위해 현재 버전에서는 ICL만이 행위 인지를 담당한다. ICL은 신호 처리와 기계 학습 기법을 사용하여 행위 인지 모델을 생성하며, 기본적인 일상 생활 행위를 구분한다[17]. 주요 입력 데이터는 신체 움직임 데이터로 스마트폰 및 스마트워치로부터 가속도 데이터를 수집한다. 수집된 데이터는 오버랩되지 않은 3초간의 데이터를 하나의 윈도우로 관리하며, 시간 및 주파수 특징을 추출하여 가우시안 혼합 모델을 통해 행위를 분류한다.

건강 및 웰니스 지식은 의료 전문가에 의해 생성되며 이는

KCL의 지식베이스에 저장된다. 이를 위해 간단한 룰 저작 툴 [18]을 이용하여 룰을 생성하며, 기반 증거자료 및 용어들은 신체 관리에 대한 정의와 행위 증진 계획을 위해 사용된다[19].

SCL은 DCL, ICL, KCL이 생성한 내용을 바탕으로 개인화된 행위 추천 내용을 생성한다. 서비스 오케스트레이터에 의해 요청이 발생하면, 보유한 지식과 사용자 데이터를 기반으로 룰 기반 추론[20]을 통해 추천 정보를 생성한다. 사용자의 건강 및 웰니스 데이터는 서비스 카탈로그의 보조 서비스를 통하여 적절한 입력 쿼리로 변경되며, 사용자의 이상적인 몸무게 및 하루 칼로리 소모량 등을 설정하는데 사용된다. 추론 과정에서 인터프리터는 지식 베이스 내의 각 룰을 분석하고 알맞은 룰을 순방향 추론 (Forward Chaining) 기법을 통해 추론을 진행한다[21]. 추천 정보는 콘텐츠 기반 여과법 (Contents-based filtration)[22]을 사용하여 사용자 개인의 행위 수준과 선호하는 행위 정보를 바탕으로 알맞은 행위 정보를 추천한다.

SL의 보안 및 프라이버시 컴포넌트는 모든 계층에 분산되어 동작한다. 사용자의 의료 및 건강 데이터에 대한 보관 및 처리는 미국의 HIPAA법을 준수하며 AES, private matching [23], 익명화[24] 등의 암호화 기법을 적용하여 데이터의 노출을 방지한다. 시스템은 퍼블릭 클라우드에 배치되기 때문에 정확도의 손실 없이 직접적인 암호화 처리를 할 수 있는 방안이 요구되며, 이를 위해 oblivious term matching[25] 기술을 사용한다.

V. 사례연구

1. 어플리케이션

마이닝 마인즈 프레임워크의 실제 사용을 위하여 안드로이드용 어플리케이션이 개발되었다. 이 앱은 개인의 체중 관리 지원을 목적으로 하며, 사용자의 신체 활동 행위를 분석하고 그에 따른 알맞은 행위를 추천한다. 기존의 건강 관리 앱과는 달리 마이닝 마인즈 앱은 최종 사용자 인터페이스 역할을 하며, 핵심 동작은 클라우드에 배치된 마이닝 마인즈 프레임워크가 담당한다. 이는 스마트폰 자원과 저장 용량의 절약, 연산량과 배터리 소모량을 줄일 수 있는 장점을 가지고 있으며 번거로운 앱의 업데이트가 불필요하고 서로 다른 단말기에서도 동일한 정보를 공유하기가 쉽다.

〈그림 2〉는 앱의 주요 화면을 나타내며, 사용자의 행위 상태, 현재 및 목표 체중, 하루 필요 소모 칼로리를 표시한다(〈그림 2.a〉). 모든 관련 연산은 사용자 프로파일 정보와 임상 전문가의 지식에 의해서 마이닝 마인즈 프레임워크에서 이루어진



그림 2. 마이닝 마인즈 어플리케이션 화면

다. 앱 화면에서는 사용자의 행위 패턴을 분석하여 최근 사용자가 소모한 칼로리량을 표시한다. 사용자가 휴대하고 다니는 스마트폰의 가속도 센서 데이터를 이용하여 사용자의 행위를 측정하며, 데이터는 수집 즉시 LTE 혹은 와이파이를 통해 마이닝 마인즈 프레임워크로 실시간 스트리밍 데이터 형태로 전송된다. 사용자의 목표 체중 도달을 돕기 위해 시스템은 사용자에게 필요한 운동 추천 정보를 알기 쉽게 표기한다. 추천 정보에는 정확한 소요 시간, 행하는 방법, 해야만 하는 이유 등을 보여 주어 사용자에게 동기를 부여하며, 이러한 방법은 사용자의 프로파일 정보를 토대로 각 개인에게 다르게 부여된다. 앱은 사용자의 몸무게 변동 추이를 보여주거나(〈그림 2.b〉), 소모한 칼로리별 행위 패턴(〈그림 2.c〉), 그리고 달력을 통한 월별 행위 수행 내역을 보여준다(〈그림 2.d〉). 제일 마지막 그림인 〈그림 2.d〉는 라이프로그 컨셉의 간소화 버전으로 볼 수 있으며 사용자의 자가 관리 및 조절을 도와준다.

2. 한계점

마이닝 마인즈는 곧 다가올 IOT 시대를 겨냥한 프레임워크로, 전제조건은 항상 통신이 가능한 상태여야 한다. 현재 마이닝 마

인즈 앱은 와이파이나 LTE 접속 상태에서만 동작한다. 와이파이 연결은 사용자들의 통신비에 대한 부담을 덜어줄 수 있지만 와이파이가 설치된 특정 공간에서만 사용할 수 있다는 단점이 있다. 또한 공간의 제약이 없는 LTE 접속의 경우 방대한 데이터 전송량으로 인해 부담이 될 수 있다. 마이닝 마인즈 앱의 통신량을 측정해 본 결과 분당 500k의 센서 데이터를 전송하며, 이는 한달 약 30기가의 용량을 사용하는 것이다. 여기에 현재의 가속도 데이터만이 아닌 더 다양한 센서 데이터를 합친다면 대용량 데이터가 전송될 것이다. 따라서 향후 5G가 현실화되어 현재의 LTE 속도를 유지할 수 있는 데이터 무제한 정책 요금제가 도입되거나, 압축 센싱 기술 [26]을 이용하여 데이터 전송의 효율성을 높이는 방법을 고려해야 한다. 압축 센싱 기술 이외에 통신 비용을 줄이는 방법은 단말 기기의 배터리 소모를 줄이는 방향으로도 사용될 수 있는 장점이 있다. 예를 들어 사용자가 아무런 동작을 하지 않을 시 데이터 전송을 일시중단 하는 방법 등이 있다.

VI. 결론

본 고에서는 개인화된 건강 및 웰니스 서비스를 제공하는 혁신적인 마이닝 마인즈 프레임워크를 소개하였다. 본 프레임워크는 최신의 디지털 헬스 패러다임에 필요한 요구사항을 반영하여 설계되었다. 또한 독자적인 아키텍처를 설계하여 개인화된 건강 및 웰빙 지원을 위한 큐레이션과 데이터, 정보, 지식 및 서비스의 마이닝을 구현하였다. 마이닝 마인즈 프로토타입을 구현하여 실제 성능을 테스트하고 발전 가능성을 입증하였다. 그리고, 향후 아키텍처의 확장 및 향상, 새로운 컴포넌트 추가 및 신기술을 도입하여 스마트한 개인화 건강 및 웰니스 서비스 제공을 위한 플랫폼으로 더욱 발전시켜 나갈 예정이다.

Acknowledgement

본 고는 산업통상자원부 산업핵심기술개발사업으로 지원된 연구결과입니다 [10049079, 퍼스널 빅데이터를 활용한 마이닝 마인즈 핵심 기술 개발].

참고 문헌

- [1] T. Hafner and J. Shiffman, "The emergence of global attention to health systems strengthening," *Health Policy and Planning*, vol. 28, no. 1, pp.41–50, 2013.
- [2] L. Hood and M. Flores, "A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory," *New Biotechnology*, vol. 29, no. 6, pp. 613–624, 2012.
- [3] E. M. Matheson, D. E. King, and C. J. Everett, "Healthy lifestyle habits and mortality in overweight and obese individuals," *The Journal of the American Board of Family Medicine*, vol. 25, no. 1, pp. 9–15, 2012.
- [4] M. M. Gillen, C. N. Markey, and P. M. Markey, "An examination of dieting behaviors among adults: Links with depression," *Eating Behaviors*, vol. 13, no. 2, pp. 88–93, 2012.
- [5] W. Demark-Wahnefried and L. W. Jones, "Promoting a healthy lifestyle among cancer survivors," *Hematology/Oncology Clinics of North America*, vol. 22, no. 2, pp. 319–342, 2008.
- [6] M. Swan, "Health 2050: the realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen," *Journal of Personalized Medicine*, vol. 2, no. 3, pp. 93–118, 2012.
- [7] A. C. Powell, A. B. Landman, and D. W. Bates, "In search of a few good apps," *The Journal of the American Medical Association*, vol. 311, no. 18, pp. 1851–1852, 2014.
- [8] O. Banos, C. Villalonga, M. Damas, P. Gloesekoetter, H. Pomares, and I. Rojas, "Physiodroid: Combining wearable health sensors and mobile devices for a ubiquitous, continuous, and personal monitoring," *The Scientific World Journal*, vol. 2014, no. 490824, pp. 1–11, 2014.
- [9] J. Oresko, Z. Jin, J. Cheng, S. Huang, Y. Sun, H. Duschl, and A. C. Cheng, "A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 734–740, May 2010.

- [10] H. Hermens, H. op den Akker, M. Tabak, J. Wijsman, and M. Vollenbroek, "Personalized coaching systems to support healthy behavior in people with chronic conditions," *Journal of Electromyography and Kinesiology*, vol. 24, no. 6, pp. 815–826, 2014.
- [11] P.-H. Chen and H.-M. Chen, "Framework design—integrating an android open platform with multi-interface biomedical modules for physiological measurement," *Journal of Convergence Information Technology*, vol. 7, no. 12, pp. 310–319, 2012.
- [12] G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, and R. Jafari, "Enabling effective programming and flexible management of efficient body sensor network applications," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 1, pp. 115–133, January 2013.
- [13] A. Gaggioli, G. Pioggia, G. Tartarisco, G. Baldus, D. Corda, P. Cipresso, and G. Riva, "A mobile data collection platform for mental health research," *Personal Ubiquitous Computer*, vol. 17, no. 2, pp. 241–251, 2013.
- [14] G. Yang, L. Xie, M. Mantysalo, X. Zhou, Z. Pang, L. D. Xu, S. Kao-Walter, Q. Chen, and L.-R. Zheng, "A health-iot platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2180–2191, Nov 2014.
- [15] O. Banos, R. Garcia, J. A. Holgado, M. Damas, H. Pomares, I. Rojas, A. Saez, and C. Villalonga, "mHealthDroid: a novel framework for agile development of mobile health applications," in *Proceedings of the 6th International Work-conference on Ambient Assisted Living and Active Ageing*, 2014.
- [16] D. Estrin and I. Sim, "Open mhealth architecture: An engine for health care innovation," *Science*, vol. 330, no. 6005, pp. 759–760, 2010.
- [17] M. Han, L. T. Vinh, Y.-K. Lee, and S. Lee, "Comprehensive context recognizer based on multimodal sensors in a smartphone," *Sensors*, vol. 12, no. 9, pp. 12 588–12 605, 2012.
- [18] T. Ali, M. Hussain, W. Ali Khan, M. Afzal, and S. Lee, "Authoring tool: Acquiring sharable knowledge for smart CDSS," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, pp. 1278–1281.
- [19] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett, C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, and A. S. Leon, "2011 compendium of physical activities: a second update of codes and met values," *Medicine and Science in Sports and Exercise*, vol. 43, no. 8, pp. 1575–1581, 2011.
- [20] C. L. Forgy, "Rete: A fast algorithm for the many pattern/many object pattern match problem," *Artificial intelligence*, vol. 19, no. 1, pp. 17–37, 1982.
- [21] J.-F. Baget, "Improving the forward chaining algorithm for conceptual graphs rules," in *International Conference on Principles of Knowledge Representation and Reasoning*, 2004, pp. 407–414.
- [22] S. Gong, "Learning user interest model for content-based filtering in personalized recommendation system," *International Journal of Digital Content Technology and its Applications*, vol. 6, no. 11, pp. 155–162, 2012.
- [23] S. Zhong and T. Chen, "An efficient identity-based protocol for private matching," *International Journal of Communication Systems*, vol. 24, no. 4, pp. 543–552, 2011.
- [24] G. Ghinita, P. Karras, P. Kalnis, and N. Mamoulis, "A framework for efficient data anonymization under privacy and accuracy constraints," *ACM Transactions on Database Systems*, vol. 34, no. 2, p. 9, 2009.
- [25] Z. Pervez, A. A. Awan, A. M. Khattak, S. Lee, and E.-N. Huh, "Privacyaware searching with oblivious term matching for cloud storage," *The Journal of Supercomputing*, vol. 63, no. 2, pp. 538–560, 2013.
- [26] S. Li, L. D. Xu, and X. Wang, "A continuous biomedical signal acquisition system based on compressed sensing in body sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1764–1771, Aug 2013.

약 력



Oresti Baños

2009년 University of Granada 석사
 2014년 University of Granada 박사
 2011년 Swiss Federal Institute of Technology Zurich 연구원
 2014년~현재 경희대학교 컴퓨터공학과 박사 후 연구원
 관심분야 : 웨어러블 센서, 모바일 컴퓨팅, 기계학습, 패턴인식



Wajahat Ali Khan

2009년 National University of Sciences and Technology 연구원
 2014년 경희대학교 컴퓨터공학과 박사
 2014년~현재 경희대학교 컴퓨터공학과 박사 후 연구원
 관심분야 : 온톨로지, SOA, 시멘틱 웹 서비스



Muhammad Bilal Amin

1998년 MAJU Islamabad Campus 연구원
 2008년 Direct Energy Application Architect
 2015년 경희대학교 컴퓨터공학과 박사
 2015년~현재 경희대학교 컴퓨터공학과 박사 후 연구원
 관심분야 : 클라우드 컴퓨팅, HPC (High performance Computing)



허 태 호

2010년 경희대학교 컴퓨터공학과 학사
 2010년~현재 경희대학교 컴퓨터공학과 석박사통합과정
 관심분야 : 유비쿼터스 컴퓨팅, 행위인지, 상황인지



방 재 훈

2007년 평택대학교 디지털응용정보학과 학사
 2013년 경희대학교 컴퓨터공학과 석사
 2013년~현재 경희대학교 컴퓨터공학과 박사과정
 관심분야 : 유비쿼터스 컴퓨팅, 모바일기반 감정인식

약 력



강 동 옥

2014년 경희대학교 컴퓨터공학과 학사
 2015년~현재 경희대학교 컴퓨터공학과 석사과정
 관심분야 : 유비쿼터스 컴퓨팅, 기계학습



Maqbool Hussain

2006년 Kohat University of Sciences and Technology 학사
 2008년 National University of Sciences and Technology 석사
 2011년~현재 경희대학교 컴퓨터공학과 박사과정
 관심분야 : 임상 의사결정 지원 시스템(CDSS), 헬스케어 시스템



Muhammad Afzal

2009년 National University of Sciences and Technology 석사
 2012년~ 현재 경희대학교 컴퓨터공학과 박사과정
 관심분야 : 임상 의사결정 지원 시스템(CDSS), 헬스케어 시스템



Taqdir Ali

2006년 Kohat University of Sciences and Technology 학사
 2011년 Interior Ministry Hail 선임 소프트웨어 엔지니어
 2012년~현재 경희대학교 컴퓨터공학과 석박사통합과정
 관심분야 : 임상 의사결정 지원 시스템(CDSS)



이 승 룡

1978년 고려대학교 재료공학과 학사
 1987년 Illinois Institute of Technology, 전산학과 석사
 1991년 Illinois Institute of Technology 전산학과 박사
 1992년 Governors State University, USA, 조교수
 1993년~현재 경희대학교 컴퓨터공학과 교수
 관심분야 : 유비쿼터스 컴퓨팅, 상황인지, 실시간 시스템, 미들웨어 시스템, 클라우드 컴퓨팅, 헬스케어시스템