

Bio-Sensing Convergence Big Data Computing Architecture

Myung-Sook Ko[†] · Tae-Gyu Lee^{**}

ABSTRACT

Biometric information computing is greatly influencing both a computing system and Big-data system based on the bio-information system that combines bio-signal sensors and bio-information processing. Unlike conventional data formats such as text, images, and videos, biometric information is represented by text-based values that give meaning to a bio-signal, important event moments are stored in an image format, a complex data format such as a video format is constructed for data prediction and analysis through time series analysis. Such a complex data structure may be separately requested by text, image, video format depending on characteristics of data required by individual biometric information application services, or may request complex data formats simultaneously depending on the situation. Since previous bio-information processing computing systems depend on conventional computing component, computing structure, and data processing method, they have many inefficiencies in terms of data processing performance, transmission capability, storage efficiency, and system safety. In this study, we propose an improved biosensing converged big data computing architecture to build a platform that supports biometric information processing computing effectively. The proposed architecture effectively supports data storage and transmission efficiency, computing performance, and system stability. And, it can lay the foundation for system implementation and biometric information service optimization optimized for future biometric information computing.

Keywords : Bio-Sensing, Wearable Device, Digital Patch, Big Data, Hadoop, Dynamic Process

바이오센싱 융합 빅데이터 컴퓨팅 아키텍처

고 명 숙[†] · 이 태 규^{**}

요 약

생체정보 컴퓨팅은 생체신호 센서와 컴퓨터 정보처리를 융합한 정보시스템에 기초하여 컴퓨팅시스템 뿐만 아니라 빅데이터 시스템에 크게 영향을 미치고 있다. 이러한 생체정보는 지금까지의 텍스트, 이미지, 동영상 등의 전통적인 데이터 형식과는 달리 생체신호의 의미를 부여하는 값은 텍스트 기반으로 표현되고, 중요한 이벤트 순간은 이미지 형식으로 저장하며, 시계열 분석을 통한 데이터 변화 예측 및 분석을 위해서는 동영상 형식 등 비정형데이터를 포함하는 복합적인 데이터 형식을 구성한다. 이러한 복합적인 데이터 구성은 개별 생체정보 응용서비스에서 요구하는 데이터의 특징에 따라 텍스트, 이미지, 영상 형식 등으로 각각 분리되어 요청되거나, 상황에 따라 복잡 데이터를 동시에 요구할 수 있다. 기존 생체정보 컴퓨팅 시스템들은 전통적인 컴퓨팅 구성요소, 컴퓨팅 구조, 데이터 처리 방법 등에 의존하므로 데이터 처리성능, 전송능력, 저장효율성, 시스템안전성 등의 측면에서 많은 비효율성을 내포하고 있다. 본 연구에서는 생체정보 처리 컴퓨팅을 효과적으로 지원하는 생체정보 빅데이터 플랫폼을 구축하기 위해 개선된 바이오센싱 융합 빅데이터 컴퓨팅 아키텍처를 제안한다. 제안 아키텍처는 생체신호관련 데이터의 저장 및 전송 효율성, 컴퓨팅 성능, 시스템 안정성 등을 효과적으로 지원하며, 향후 생체정보 컴퓨팅에 최적화된 시스템 구현 및 생체정보 서비스 구축을 위한 기반을 제공할 수 있다.

키워드 : 바이오센싱, 웨어러블 디바이스, 디지털 패치, 빅데이터, 하둡, 동적 프로세스

1. 서 론

고령화 사회가 본격화되고 웰빙에 대한 높은 사회적 관심

이 스마트폰의 보편화와 맞물리면서 생체신호를 지속적으로 모니터링 해주는 웨어러블컴퓨터에 대한 수요가 증가하고 있다. 웨어러블디바이스는 웨어러블컴퓨터 또는 착용컴퓨터로도 불리는데 안경, 시계, 밴드, 의복 등과 같이 착용할 수 있는 형태로 된 컴퓨터를 의미한다. 웨어러블디바이스는 사용자가 거부감 없이 신체의 일부처럼 항상 착용하고 사용할 수 있으며 신체에 직접 접촉하여 각종 정보를 수집하고 정보화할 수 있는 차세대 정보디바이스이다[1-3].

* 이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2017TR1D1A3B04033897).

† 정 회 원 : 부천대학교 경영과 교수

** 중 심 회 원 : 원광대학교 바이오나노화학부 교수

Manuscript Received : October 10, 2017

Accepted : October 27, 2017

* Corresponding Author : Tae-Gyu Lee(tglee@wku.ac.kr)

웨어러블디바이스는 유형에 따라 휴대형, 착용형, 부착형 등으로 분류할 수 있다. 휴대형의 경우 스마트폰과 연계하여 작동하는 제품으로 구글글래스, 갤럭시기어, 애플워치, 스마트의류가 있다. 착용형은 웨어러블디바이스의 최종 단계로 전망되는 형태로 인체에 일정시간 머물다 자연적으로 분해되어 없어지는 형태로도 활발하게 연구개발되고 있다[2]. 그리고 피부를 통한 물질 전달이나 생체신호 센싱 기술이 발달함에 따라 피부를 활용한 의료기술이 많은 발전을 거듭하고 있으며 특히, 탈부착이 쉽고 사용이 간편한 패치타입의 부착형 웨어러블디바이스가 개발됨에 따라 다양한 디지털패치의 사용이 기대된다[3]. 패치나 패브릭 소재를 기반으로 한 웨어러블컴퓨터는 섬유 소재 기술과 의료기기 융합 연구를 통해 지속적으로 증가할 전망이다[2].

스마트폰 시대 이전의 웨어러블 컴퓨팅은 PC나 노트북을 소형화하는데 초점이 맞추어져 있었으며 다른 한편으로는 센서를 통해 스마트 기기와 연동하는 형태의 디바이스들도 출시되기 시작했다. 웨어러블디바이스의 장점은 주변 환경에 대한 정보나 신체 변화를 실시간으로 끊이지 않고 지속적으로 수집할 수 있다는 것이다[4]. 최근 스마트 모바일 환경에서 웨어러블디바이스는 의료서비스 패러다임의 변화를 주도할 것으로 예상된다.

그러나 실시간 모니터링을 통해 사용자 건강상태가 기록된다면 데이터양이 방대할 수밖에 없다[2]. 웨어러블디바이스는 단순히 기기의 소형화뿐만 아니라 디바이스로서의 IoT 단말기이자 신체정보 수집 후 서버로 전송하여 의료건강 빅데이터(Big Data)를 구축하는 단말기로서 건강 및 운동 관리, 핀테크 등 다양한 정보관련 산업에 연관되어 있다[4]. 따라서 수집된 빅데이터를 서버에 전달하고 처리 및 분석하는 방법에 대한 연구가 강화되어야 한다. 의료데이터의 실험적 분석은 많은 인과 관계를 규명하며, 원인-결과 상관계는 확실적인 분포를 가지고 수집된 데이터의 여러 특징들에 의해 특정그룹으로 분류하여 효과적인 진단 및 처방에 도움을 줄 수 있다[7]. 이러한 생체정보 데이터의 다양한 분석을 지원하는 데이터마이닝 기술은 인체의 임상적인 특징을 다수 포함하고 있는 의료데이터에 숨겨진 지식을 발견하여 구조화시킴으로써 새로운 의료지식을 창출할 수 있다[8-10]. 따라서 실시간으로 수집되는 다양한 형태의 센싱 데이터 수집방법과 더불어, 수집된 방대한양의 데이터처리를 지원하는 빅데이터 융합 아키텍처에 대한 연구가 필요하다.

본 연구는 바이오센싱 융합 빅데이터 컴퓨팅을 효과적으로 지원하기 위한 플랫폼을 구축하기 위해 개선된 아키텍처를 제안한다. 이러한 제안 아키텍처는 데이터의 저장 및 전송 효율성, 컴퓨팅 성능, 시스템 안정성 등을 효과적으로 지원한다. 특히, 향후 생체정보 컴퓨팅에 최적화된 시스템 구현 및 생체정보서비스 구축을 위한 기반을 제공하고자 한다. 제 2장에서는 관련 연구에 대해 다루고 제 3장에서는

실시간 바이오센싱 융합 아키텍처 및 바이오센싱 빅데이터 융합 컴퓨팅 프로세스에 대한 연구방법을 기술한다. 제 4장에서는 제안한 아키텍처 및 프로세스에 대한 성능을 비교 분석하고 제 5장에서 결론 및 향후 연구 과제에 대하여 기술한다.

2. 관련 연구

기존 웨어러블디바이스의 응용서비스 데이터 수집을 위해서 응용단말마다 복잡한 사용자설정 등의 한계점을 극복하고, 단말에서 수집한 정보를 이웃 단말 간의 자율적 협업을 의해 데이터를 수집할 수 있는 프로토콜을 제안하였다[5].

생체정보시스템의 센싱정보처리는 생체신호 센싱, 데이터 변환, 데이터 필터링, 데이터 전송, 데이터 저장, 데이터 분석 및 센싱 장치의 통계로 구성된다. 기존 생체정보시스템의 구조와 프로세스의 특징은 다음과 같다. [14]는 멀티 에이전트 기법을 포함한 개별 기능을 지원하는 에이전트기법을 지원하여 시스템구조 및 운영능력을 최적화시키는 특성을 보여주기 때문에 생체정보에 특화된 개별프로세스의 성능을 향상시킨다. [15]는 다양한 생체정보 워크로드를 지원하는 데이터마이닝의 구조적인 문제를 기술하였다. 특히 실행시간, 확장성 및 병목조건을 고려한 메모리 계층구조의 특성을 분석하였다. [16]은 생체신호 데이터 처리를 위해 대규모 프레임워크, 분산컴퓨팅 엔진, 순차분석 모듈, 마이크로레이 분석 모듈 등을 지원하는 소프트웨어 시스템을 구현할 필요성을 제시하였다.

생체정보는 다양한 수준에서 거대한 정보의 중요한 원천이다[16]. 모든 프로세스의 디지털화 및 고효율 장치의 가용성이 낮은 비용으로, 생체정보 연구를 비롯한 모든 곳에서 데이터 볼륨이 증가하고 있다. 예를 들어, 단일 서열화 인간 게놈의 크기는 약 200기가바이트이다[17]. 유럽 생체정보 연구소 (European Bioinformatics Institute, EBI)는 2013년에 비해 22페타바이트와 매년 총 저장용량이 두 배가되는 반면, 2014년에는 유전자, 단백질 및 소분자에 대해 약 40페타바이트의 데이터를 보유하고 있다[18].

[17]은 생체정보 프로세스의 지식 발견을 위한 6단계를 구성하였다. 이러한 분석 프로세스는 시스템 요구사항 또는 환경에 따라 선택적이고 적응적으로 사용된다. 이러한 분석 방법은 분석서버가 있는 글로벌스테이션이나 오픈스테이션에서 제한적으로 사용할 수 있으며 사용자접근 방식, 성능 스케줄링, 로드밸런싱, QoS 및 내결함성을 고려한 아키텍처의 필요성이 제시되었다.

생체정보는 개인마다 더 많은 생체요소 데이터를 생성할 수 있는 기술로 인해 지속적으로 증가하고 있다. 생체정보 분야에서 데이터의 양이 증가함에 따라 빅데이터를 효율적으로 처리할 수 있는 데이터마이닝 및 분석기술 개발의 중

요성이 대폭 증가되었고, 빅데이터는 생체정보와 같이 매우 민감한 분야의 연구에 유용하다. 생체정보의 주요 쟁점은 스토리지, 데이터전송, 보안 및 개인정보보호의 근원에 대한 데이터처리에 관한 것이다. [11-13]은 다층 프로세스 시나리오를 고려하지 않고 동적 프로세스 구성과 단일 프로세스 시나리오를 보여주었다. 기존의 생체정보처리 시스템 구성의 초기 단계에서 기본 및 선택적 프로세스로 정의된 후 변경 없이 단일 정적 및 동일 프로세스를 수행한다.

그러나 이러한 정적 프로세스는 모바일컴퓨팅을 수행하는 모바일 생체정보 시스템의 응용에서 비효율적인 실행을 나타내며 특히, 생체정보시스템의 프로세스 구성 과정에서 새로운 정의 및 실행의 초기화를 수행하는 불편한 방법이 수반된다. 특히, 기존 생체정보시스템은 비정형 생체데이터를 정형 데이터로 변환하여 처리하는데 집중하였고 정형화된 데이터의 의미 분석에 초점을 맞추어왔다. 이러한 접근 방법은 불필요한 변환과정을 유도한 동시에 비정형 생체데이터의 의미분석 오류 한계성을 내포하고 있다.

본 연구의 빅데이터 생체정보 플랫폼 아키텍처는 생체정보의 정형 및 비정형 데이터를 통합적으로 운영할 수 있는 생체정보 다층 구조 및 동적 프로세싱 방법을 제안한다. 이러한 플랫폼 구조는 비정형 생체데이터에 대한 비정형 의미 분석을 위한 다층 동적 프로세스 설계 및 실행 방법을 제안한다. 본 논문에서는 기본적인 웨어러블디바이스 생체 정보 센싱 아키텍처와 동적 프로세스 설계를 위한 빅데이터 기반 융합 동적 프로세스 구성 방법에 대해 기술한다.

3. 연구 방법

3.1 바이오 센싱을 통한 정보 수집

본 연구에서 제안하는 디지털 생체정보시스템의 구조인 바이오융합 정보 아키텍처는 다음 Fig. 1과 같다. 그림의 아키텍처는 웨어러블디바이스 착용자 또는 환자의 병원 디지털침대매트를 병원 또는 회복시설에서 사용할 수 있는 생체정보를 수집하는 구성이다. 모바일건강관리 구성요소로서, 멀티 바이오센서를 갖춘 착용형 클라이언트는 모바일 건강관리를 위한 바이오센싱 아키텍처에 통합될 수 있다.

생체정보 센싱시스템 구조는 개인용 생체정보 장치를 통해 각 사용자의 생체정보를 수집한다. LAN 및 PSTN과 같은 유선 네트워크를 사용하거나 WIFI, Zigbee 및 Bluetooth와 같은 무선 네트워크를 사용하여 로컬 인프라 네트워크를 구성한 후 이러한 생체정보를 수집한다. 뿐만 아니라 유선 IP 네트워크 또는 무선 3G/4G 글로벌네트워크를 통해 생체 데이터를 전송할 수 있다. 또한 개인 사용자의 생체 정보를 모니터링하고 관리하는 역할은 인프라네트워크에 위치한 개인용 컴퓨터 또는 생체정보 서버를 통해 수행된다[14]. 이러한 모바일 생체정보 기기를 통해 수집된 생체정보는 센싱방

식, 전송채널, 저장 및 필터링 과정, 분석 및 평가 방법 등에 따라 서로 다른 정보처리 시스템을 구성한다[15, 16].

생체정보 빅데이터 시스템 구조는 Fig. 1과 같이 3계층으로 센싱부, 전송부, 관리부 등으로 구성된다.

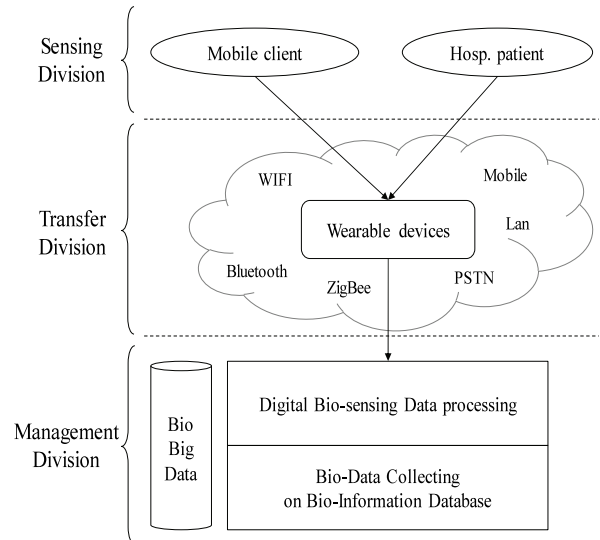


Fig. 1. Bio-information Big Data System Architecture

첫째, 센싱부 층(Sensing Division Layer)은 클라이언트가 웨어러블컴퓨팅 환경에서 생체정보를 생성하며, 컴퓨팅시스템은 생체정보와 상호작용하는 사용자에게 다양한 수집 및 분배 인터페이스를 제공한다. 또한 다양한 건강가이드서비스와 의료정보서비스를 사용자에게 제공할 수 있다. 인터넷상의 사용자는 서버에서 제공되는 표준 공유정보를 수신하며 개방된 의료정보를 다른 사람들과 공유하기 위해 사용자의 요구사항과 관련 추가 정보를 서버에 업로드 할 수 있다. 사용자에게 대한 건강분석 정보의 디스플레이 인터페이스는 위험 및 경고 신호 이벤트, 실시간 건강정보 분석 이벤트, 호흡 및 건강관련 신호 이벤트, 체격상태 분석정보의 표시 이벤트가 포함되며, 생체정보의 제어 및 모니터링에 필요한 생체정보 입출력 인터페이스가 정의된다.

둘째, 전송부 층(Transfer Division Layer)은 사용자와 서버 간의 인트라넷과 인터넷의 네트워킹 계층인 투명 프로세스 구성 요소로서, 수집된 정보를 사용자로부터 서버로 전달하거나 또는 피드백 제어 및 정보를 서버에서 사용자에게로 전달한다. 이 계층은 등록된 사용자의 상태정보를 보호하기 위한 보안인증 서비스를 기반으로 한다. 개방형 투명 레이어는 서버에서 받은 표준 공유 상태정보를 인터넷상의 사용자에게 공개하며, 인터넷 사용자의 요구사항을 서버에 전달한다.

셋째, 관리부 층(Management Division Layer)은 건강관리센터의 서버에서 수집된 생체정보를 저장하고, 수집된 정보에 대한 평가 및 분석을 수행하고, 평가된 결과에 따라

피드백제어 및 고객센터를 제공한다. 또한, 인터넷에 공개된 생체정보서비스를 제공하기 위한 표준 건강정보를 작성하며 사용자 생체신호에 기반 사용자건강상태 모델을 정의한다. 임계값수 및 임계값은 만족/경고/위험 등을 분류하기 위해 생체정보기기의 제약관리 프로세스를 정의한다. 디지털 장치의 식별시스템 및 상태정보 관리를 위하여 개별 사용자의 단위 통계분석 및 평가모델을 정의한다.

3.2 바이오센싱 빅데이터 융합 처리 컴퓨팅 아키텍처

본 연구에서 제안하는 바이오센싱 빅데이터 융합 처리 컴퓨팅 아키텍처는 Fig. 1과 같이 생체정보를 수집하는 디지털 센싱부, 수집된 생체정보를 사용자 단말이나 관리서버로 전송하는 생체정보 전송부, 그리고 수신한 바이오 정보를 저장, 분석 및 피드백하는 바이오정보 관리부 등의 계층구조로 이루어져 있다.

클라우드 컴퓨팅의 대중화는 하둡 및 관련 소프트웨어와 같은 구현을 지원하는 오픈 소스 소프트웨어에 어느 정도 기인한다. 하둡은 여러 분산컴퓨팅 노드에서 대규모 데이터 세트를 분산처리할 수 있으며 HDFS, HBase와 같은 대규모 데이터 스케일링을 지원하고 무정지 병렬 처리 분석을 가능하게 한다. 따라서 하둡은 데이터 크기를 특정할 수 없는 실시간 비정형 데이터집합으로 구성되는 생체정보컴퓨팅의 요구를 충족시킬 수 있다. 몇몇 연구는 하둡을 생체정보시스템에 성공적으로 적용하여 클라우드 기반 생체데이터관리 자원으로 이끌었다[9-11]. 클라우드컴퓨팅은 인터넷을 통해 분산컴퓨팅 서비스를 제공하며, 생체정보 클라우드는 센싱 데이터 수집, 생체데이터 저장, 생체정보 분석 등에 이르기까지 다양한 생체정보컴퓨팅 서비스를 포함한다. 생체정보 클라우드 기반 서비스는 다음 Fig. 2와 같이 DaaS(Data as a Service), SaaS(Software as a Service), PaaS(Platform as a Service), IaaS(Infrastructure as a Service)로 분류되는 클라우드 계층모델 기반 생체정보 컴퓨팅 플랫폼을 제안한다[12].

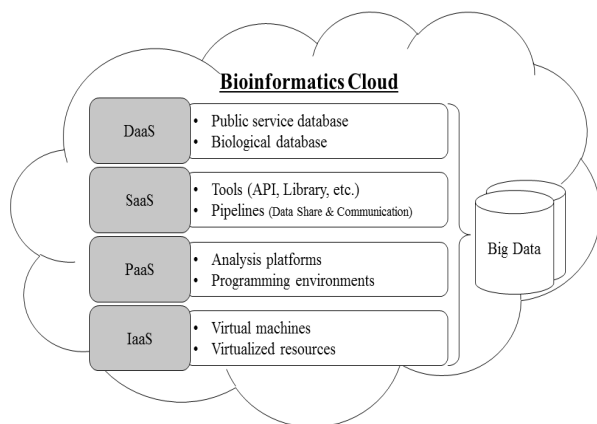


Fig. 2. Cloud-based Big Data Services for Bioinformatics

특히, 생체정보 의미분석(semantic analysis)을 위해 활용되는 데이터마이닝(data mining)은 데이터의 양적 증대와 더불어 생체데이터에 내재해있는 의미정보(meaning information)를 찾아내고 이를 분석하여 생체정보 클라이언트의 건강상태 등을 예측, 발견, 묘사, 식별 등을 목적으로 한다. 또한, 데이터마이닝은 방대한 양의 생체신호 데이터 속에서 의미패턴(meaning pattern)들을 식별해 내고 정보간의 관계정립 및 정보분석을 통하여 미래에 발생하는 생체이벤트(bio-event)를 예측하는 자동화 프로세스이다. 데이터마이닝의 역할은 최근 기술의 급속한 성장에 기반을 둔 빅데이터 시대가 도래함에 따라 그 중요성이 더욱 부각되고 있다[6].

데이터마이닝 기법은 추정-예측, 분류-예측 모델링을 수행하는 지도학습 모델링(Supervised Modeling)과 군집화, 연관규칙탐사, 연속규칙탐사 모델링을 수행하는 비지도학습 모델링(Unsupervised Modeling)으로 분류할 수 있다. 데이터마이닝을 구성하는 요소는 사전처리(data preprocessing), 데이터노이즈제거(outlier exception), 연관규칙학습(association rule training), 군집화(clustering), 분류(classification), 회귀분석(regression), 요약(summarization) 등이다.

본 연구에서는 다양한 형태로 수집된 센서 데이터를 처리하기 위해서는 대량의 자료를 처리할 수 있는 분산 컴퓨터 클러스터에서 동작하고 분산 응용프로그램을 지원하는 하둡파일시스템(Hadoop File System)을 사용한다.

하둡은 맵-리듀스(Map-Reduce) 메커니즘을 사용하는 대용량 데이터에 대한 분산처리 분석을 지원하는 프레임워크이다. MapReduce 알고리즘은 임의의 작업을 여러 작은 부분들로 나누어 분산노드에 할당하고 처리 후 결과를 수집한다. Hadoop은 대형 소스 분산 데이터분석을 위해 상용 Linux 클러스터에 설치할 수 있는 오픈소스 소프트웨어 프레임워크이다[15]. 하둡은 애플리케이션 데이터에 대한 높은 처리량의 액세스를 제공하며 대용량 데이터세트가 있는 애플리케이션에 적합하다. 또한 내결함성이 뛰어나며 저가의 하드웨어에 배치되도록 설계되었다. 하둡은 애플리케이션을 실행할 수 있는 맵-리듀스 및 하둡분산파일시스템(HDFS)으로 구성된다. 모든 노드는 맵-리듀스 및 HDFS에 기반을 둔 요구 사항에 따라 마스터 또는 슬레이브로 동작할 수 있다[16]. 하둡은 맵-리듀스와 HDFS의 두 가지 핵심모듈을 제공한다. 맵 함수(Map Method)는 키 값을 읽어 필터링하거나 다른 값으로 변환하는 작업을 수행하고 리듀스 함수(Reduce Method)는 맵 함수를 통해 출력된 결과에 대하여 새로운 키를 기준으로 그룹화하는 역할을 한다. 맵-리듀스는 컴퓨터 프로그램을 여러가지 작은 하위 문제로 나누고 여러 컴퓨터 노드에 배포하며 HDFS는 이러한 노드에 데이터를 저장하는 분산파일시스템을 제공한다. 하둡 및 관련 소프트웨어는 여러 노드간의 로드밸런싱을 처리하고 모든 노드에서 자동으로 재실행할 수 있는 노드 장애를 감지하도록 설계되었다. 하둡 프레임워크는 다음 Fig. 3과 같다.

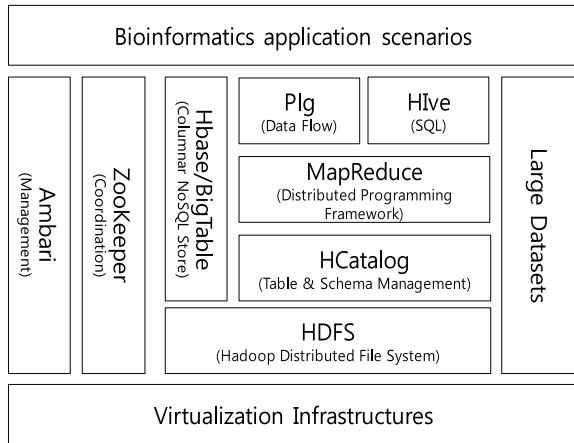


Fig. 3. Hadoop Framework for Bioinformatics

클라우드컴퓨팅은 데이터 전송의 문제를 극복하는 주요 측면 중 하나이다. 또한 데이터 처리를 위한 통합된 위치-독립 플랫폼을 제공한다. John Craig Venter Institute (JCVI), 노스웨스트 환경 사업 협의회 (NEBC), 하버드 공중 보건 대학 (Public Health of School) 등은 클라우드 컴퓨팅 플랫폼에서 BioLinux를 사용하여 계놈 분석을 가능하게 하는 클라우드 인프라를 제공하였다. BioLinux는 바이오서열 정렬 도구, 클러스터링 도구, 시퀀스 어셈블링 소프트웨어, 시각화 도구, 계놈 편집 및 계통 발생 기반 패키지를 포함하는 135 개 이상의 생체정보 소프트웨어를 제공한다[16]. 수많은 생체정보 클라우드가 존재하기 때문에 클라우드 사이의 상호 운용성과 표준화가 중요한 문제가 될 것이며 하둡 클러스터를 사용하여 새로운 전략을 개발하는 것은 생체정보에서 중요한 측면이다[18]. 다음 Fig. 4는 생체정보 빅데이터 처리

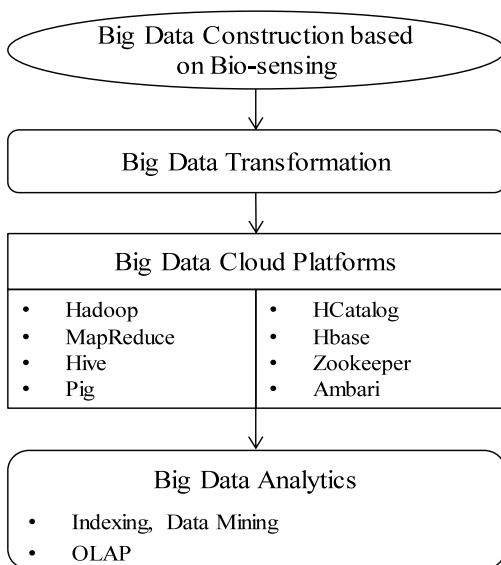


Fig. 4. Conceptual Processes for Big Data Analytics in Bioinformatics

프로세스 개념도이다.

수집된 생체정보 빅데이터는 데이터 변환 및 필터링 과정 등의 데이터 프로세싱 단계를 거친 후 생체정보 데이터베이스로 구축되며, 실시간 모니터링서버에서는 기 수집된 생체정보와 실시간으로 수집되는 정보들을 동적으로 처리할 수 있는 프로세스와 연동하여 클라이언트 사용자의 건강상태를 평가값(good/boundary/risk/others)으로 피드백해줄 수 있다. 이때, 동적할당 프로세스는 빅데이터 플랫폼을 기반으로 한다. 본 연구에서 제안하는 바이오센싱 융합 빅데이터 기반 동적 프로세스는 아래 Fig. 5와 같다.

또한 실시간으로 모니터링하고 상태 평가를 피드백하기 위해서는 빅데이터 기반의 데이터마이닝의 학습 및 분석 과정을 거쳐 데이터 분류/예측 모델을 생성하고 이를 기반으로 실시간으로 전송되는 바이오센싱 데이터에 대해서 데이터마이닝에 의해 생성된 모델을 적용하여 의미 있는 정보를 제공할 수 있는 프로세스를 제안하는 아키텍처에 추가하고자 한다.

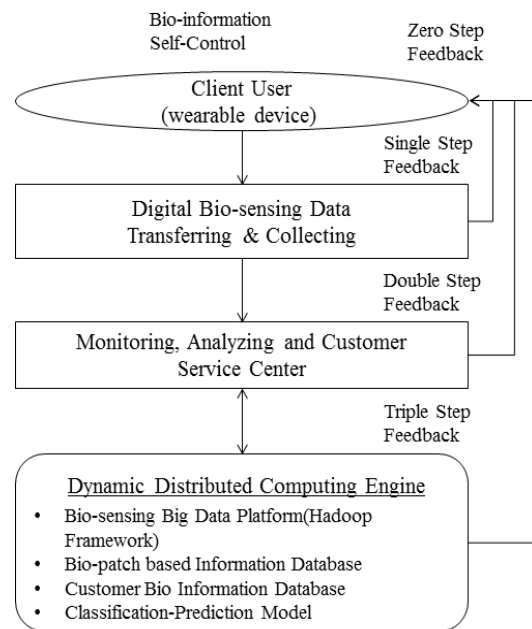


Fig. 5. Dynamic Bio-sensing and Big Data Computing Process

Fig. 5와 같이, 다층 동적 프로세스를 구축함에 있어서, 클라이언트로부터 빅데이터 분산컴퓨팅 서버까지 0단계부터 3단계까지 단계적 생체정보 프로세싱(분석 및 피드백 등)을 수행한다. 0단계에 가까울수록 실시간성이 높고 컴퓨팅 자원의 경량화 및 고성능이 요구되며, 3단계에 가까울수록 실시간성이 낮고 생체정보 컴퓨팅의 저장성 및 확장성이 요구된다.

본 연구는 실시간성의 정도에 영향을 미치는 생체정보 네

트위크 및 전송능력 구조에 따른 성능평가를 통해서, 향후 생체정보 프로세스 및 플랫폼에 대한 구성방안을 제시하고자 한다.

4. 분석 결과

제안 연구의 실험은 제안된 모바일 생체정보 융합 컴퓨팅 환경에서 필수적으로 요구되는 동적 전송 및 동적 프로세싱을 지원하는 동적 아키텍처의 성능을 분석하고자 한다.

모바일 사용자 환경에서 생체정보 센싱 빅데이터 융합 컴퓨팅 아키텍처의 동적 구성에 따라 생체정보 센싱으로부터 수집된 생체정보를 처리하는 다층 동적 프로세싱을 선택적으로 수행한다.

Table 1은 생체정보를 처리하기 위한 시간 흐름에 따른 생체정보 요구 단위 대역폭 변화에 따른 시간지연 결과를 비교한 결과이다. 결과에서 볼 수 있듯이 생체정보 센싱 Triple step이 정보의 시간지연이 Single step 생체정보의 시간지연보다 큼을 알 수 있다. 따라서 생체정보 전송대역폭이 낮은 경우는 생체정보 전송경로 상의 Step 카운트를 줄이는 것이 생체정보 전송효율을 높이는 중요한 구조적 요인으로 작용함을 확인할 수 있다.

Fig. 6은 다층 생체센싱 아키텍처에 기초한 무선전송 구간(step)에 있어서 생체정보를 처리하기 위한 생체정보 요구 단위 대역폭 변화에 따른 시간지연 변화와 관계를 나타내는 함수적 관계를 보여준다. Single step은 동적 아키텍처의 가장 단순한 형태의 구성으로 센서로부터 단말까지 단일 홉(single hop)의 구성이다. Triple step은 동적 아키텍처의 가장 일반적 형태의 구성으로 센서로부터 단말까지 세 개의 홉(triple hop)의 구성이다. 그래프에서 볼 수 있듯이 생체정보 센싱 전송대역폭이 낮은 경우가 높은 경우보다 Triple Step이 정보 전송 시간지연이 Single Step의 전송시간지연보다 영향을 더 크게 받음을 알 수 있다.

Fig. 7은 다층 생체센싱 아키텍처에 기초한 유선전송 구간(step)에 있어서 생체정보를 처리하기 위한 생체정보 요구 단위 대역폭 변화에 따른 시간지연 변화와 관계를 나타내는 함수적 관계를 보여준다. 유선전송 구간은 Fig. 6의 무선전송 구간에 비해 Single step과 Triple step의 시간지연 차이 비율은 유사하게 나타나지만, 높은 전송대역폭으로 인하여 전송시간지연이 최소화 된다.

그러므로 다층 생체신싱 아키텍처에 기반한 동적 프로세스 구성은 무선전송 매체보다는 유선전송 매체에 의존한 에이전트 컴퓨팅과 캐싱 및 버퍼링을 운영하는 전송스테이션을 구성하는 것이 시스템 안전성과 효율성을 강화시키는데 효과적이다.

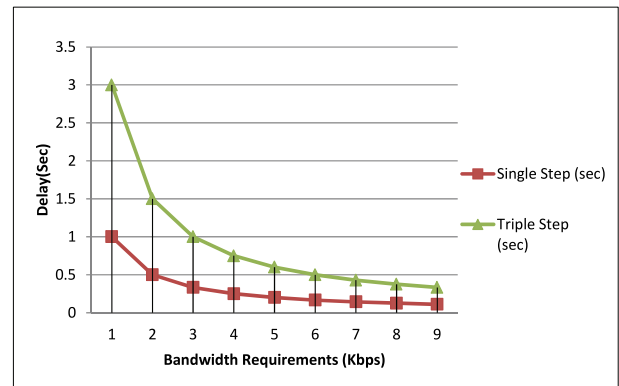


Fig. 6. Delay Analysis of Wireless Transfer in Multiple Layered Bio-sensing Architecture

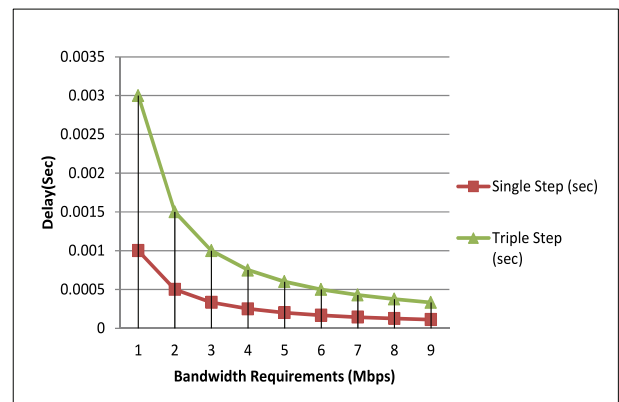


Fig. 7. Delay Analysis of Wired Transfer in Multiple Layered Bio-sensing Architecture

5. 결론

본 논문에서는 멀티 생체정보센싱시스템의 단일 정적 처리 구성에 따른 문제점을 극복하기 위해 동적 생체정보 프로세스와 멀티 센싱프로세스 방법을 제안하였다. 이는 다층

Table 1. Comparisons of Required Bandwidth for Dynamic Processing in Wireless Transfer

Bandwidth Requirements (Kbps)	1	2	3	4	5	6	7	8	9
Single Step Delay (sec)	14	12	13	14	12	13	12	13	12
Triple Step Delay (sec)	85	84	55	105	75	110	110	53	103

동적 프로세스를 구축하기 위한 확장된 다층 생체정보 센싱 구조와 동적 생체정보프로세스로 구성되어 있다. 이러한 다층 동적프로세스 구성은 시스템을 실행하는 동안 초기화 프로세스 또는 시스템 실행 중단없이 사용자 환경, 다층 생체정보 수집 유형 및 방법 등과 같은 시스템 환경의 동적 변경에 따른 적응성을 지원하는 이점을 제공한다.

동적 생체정보 센싱 프로세스 구조는 향후 멀티 생체정보 센싱시스템 플랫폼의 기능 또는 방법으로 활용될 수 있다. 특히, 생체정보시스템의 분석 및 평가를 위한 분산 빅데이터컴퓨팅 모델은 생체신호 분석 또는 생체정보의 유용성을 극대화하기 위한 미래 대안으로 고려될 수 있다. 제안하는 동적 센싱 아키텍처는 생체데이터의 저장 및 전송 효율성, 컴퓨팅 성능, 시스템 안정성 등을 효과적으로 지원할 수 있다.

References

[1] Wikipedia [Internet], <https://ko.wikipedia.org/wiki/>.

[2] M. K. Kim and H. J. Lee, "Wearable / implantable smart medical device," *JKICS*, Vol.33, No.6, pp.47-52, 2016.

[3] S. Y. Yang, K. Y. Sung, and S. G. Lim, "Research trends in medical skin patches," *BT NEWS*, The Korean Society For Biotechnology And Bioengineering, Vol.22, No.1, pp.62-68, 2015.

[4] www.daum.net [Internet], http://blog.daum.net/_blog/BlogTypeView.do?blogid=0JhIe&artid=8766959.

[5] Y. H. Oh, J. S. Lee, and S. J. Kang, "Protocol Design for Opportunistic Direct M2M Communication in Wearable Computing Environment," *JKICS*, Vol.39, No.2, pp.151-163, 2014.

[6] D. W. Lee and K. H. Kwon, "Analysis of Determinants of Labor Market Performance of Young College Students Using Data Mining Analysis Method," *Korean Policy Academic Bulletin*, Vol.25, No.3, pp.362-363, 2016.

[7] Y. G. Jung, "Feature Reduction and Bayesian Networks Learning for Medical Datamining," *Proceeding of KISSE*, Vol.31, No.1B, pp.595-597, 2004.

[8] S. Y. Han and Y. G. Jung, "A Naive Bayesian Learning of Clustering for Medical Data mining," *Proceeding of KISSE*, Vol.37, No.1C, pp.410-413, 2010.

[9] G. H. Park, "A Trends of Big Data Processing for Digital Healthcare," *The Korea Contents Association Review*, Vol.15, No.1, pp.35-37, 2017.

[10] S. S. Kim, M. C. Chung, T. W. Lee, and J. H. Won, "Implementation of Medical Data-Based Big Data Analytics Service," *Proceeding of KISSE*, pp.157-159, 2015.

[11] T. G. Lee, S. H. Lee, and H. Y. Kim, "Dynamic Configuration Method of Process Design in Bio-sensing Information Computing System," *International Journal of Bio-Science and Bio-Technology*, Vol.5, No.6, pp.147-156, 2013.

[12] T. G. Lee, "Chapter 15: Mobile Healthcare Computing in the Cloud," *Mobile Networks and Cloud Computing Convergence for Progressive Services and Applications*, IGI Global, pp.275-294, 2014. DOI: 10.4018/978-1-4666-47817.ch015

[13] T. G. Lee and S. H. Lee, "Dynamic stepping information process method in mobile bio-sensing computing environments," *Technology and Health Care*, IOS Press., 2014. DOI: 10.3233/THC-140795

[14] E. Merelli, G. Armano, N. Cannata, F. Corradini, M. d'Inverno, A. Doms, P. Lord, A. Martin, L. Milanese, S. Moeller, M. Schroeder, and M. Luck, "Agents in bioinformatics, computational and systems biology," *Briefings in Bioinformatics Advance Access*, pp.1-15, 2006.

[15] B. Ozisikyilmaz, R. Narayanan, J. Zambreno, G. Memik, and A. Choudhary, "An Architectural Characterization Study of Data Mining and Bioinformatics Workloads," In *Proceedings of the International Symposium on Workload Characterization*, 2006.

[16] A. K. Atwa, A. S. Aboelenine, M. S. Mabrouk, and Y. M. Kadah, "A New Enterprise Scale Software System for the Analysis of the Biological Data: An Enterprise Lifeware," *Proceedings of Cairo International Biomedical Engineering Conference*, pp.1-4, 2006.

[17] S. Gonzalez, V. Robles, J. M. Pena, and E. Menasalvas, "Instantiation and adaptation of CRISP-DM to Bioinformatics computational processes," *Allen Institute for Artificial Intelligence*, 2011.

[18] M. Zakarya, I. U. Rahman, N. Dilawar, and R. Sadaf, "An Integrative Study on Bioinformatics Computing Concepts, Issues and Problems," *International Journal of Computer Science Issues*, Vol.8, Issues 6, No.1, pp.330-339, 2011.



고 명 속

<http://orcid.org/0000-0002-7445-6882>

e-mail : kms@bc.ac.kr

1989년 이화여자대학교(이학사)

1993년 고려대학교 컴퓨터학과(이학석사)

1998년 고려대학교 컴퓨터학과(이학박사)

2001년~현 재 부천대학교 경영과 교수

관심분야 : 유전자알고리즘, 빅데이터경영



이 태 규

<http://orcid.org/0000-0001-7204-8132>

e-mail : tglee@wku.ac.kr

1992년 군산대학교 전자계산학과(학사)

1996년 숭실대학교 컴퓨터공학과(석사)

2006년 고려대학교 컴퓨터학과(박사)

2014년~현재 원광대학교

바이오나노화학부 교수

관심분야: Wearable Computing, Bio-Informatics, Distributed Systems, Wireless Communications, Sensor Networks