

빅데이터 기반의 가속도 신호를 이용한 집단 행동패턴 및 활동성 분석 시스템

(Group Behavior Pattern and Activity Analysis System Using Big Data Based Acceleration Signals)

김태웅*

(Tae Woong Kim)

요약

빅데이터를 이용한 데이터 분석 시스템은 정치, 교통, 자연재해, 쇼핑, 고객관리, 의료, 기상정보 등의 다양한 분야에서 활용할 가치가 있다. 특히 웨어러블 디바이스로부터 수집한 가속도 신호를 이용한 개인의 운동량 분석은 이미 보편화되어 사용되고 있다. 하지만 이러한 시스템에서 사용하는 데이터는 개인의 운동량을 측정하기에 필요한 데이터만을 저장하고 있기 때문에, 개인의 운동량외의 다양한 분석결과들은 제공하지 못하고 있다. 본 논문에서는 개인의 스마트폰에서 수집 가능한 가속도 신호를 24시간 측정하고 이를 빅데이터 기반으로 저장하여 집단 행동패턴 및 활동성 분석을 위한 시스템을 제안한다. 또한 다양한 스마트 기기에서 사용할 수 있도록 표준 메시지를 이용하여 가속도신호를 송신하고 분석결과를 수신하는 시스템을 제안한다.

■ 중심어 : 빅데이터 ; 가속도 신호 ; 활동성 분석 ; 운동량 ; 분석 시스템

Abstract

The data analysis system using Big-data is worthy to be used in various fields such as politics, traffic, natural disaster, shopping, customer management, medical care, and weather information. Particularly, the analysis of the momentum of an individual using an acceleration signal collected from a wearable device has already been widely used. However, since the data used in such a system stores only the data necessary for measuring the individual activity, it does not provide various analysis results other than the exercise amount of the individual. In this paper, I propose a system that analyzes collective behavior pattern and activity based on the acceleration signal that can be collected from personal smartphones for 24 hours a day and stored in big data. I also propose a system that sends acceleration signals and receives analysis results using standard messaging to use on various smart devices.

■ keywords : Big Data ; Acceleration Signal ; Analysis of Activity ; Amount of Exercise ; Analysis System

I. 서론

빅데이터는 최근 우리사회의 산업, 과학, 예술분야에서 공공 부문과 민간부문등 모든 영역에서 가장 주목받는 키워드이다 [1]. 인터넷, 웹, 모바일, 스마트 기기, 센서 등 정보화의 발전과 ICT 패러다임의 변화는 인류의 생활에 큰 변화를 일으키고, 더불어 데이터의 폭증을 가져오게 되었다[2].

빅데이터 기반의 서비스는 정치, 교통, 자연재해, 쇼핑, 고객 관리, 의료, 기상정보 등의 다양한 분야에서 제공되고 있다. 특히 의료분야에서는 과거와 달리 병원이 아닌 일상생활에서 생

체신호를 측정할 수 있는 웨어러블 디바이스의 등장으로 많은 종류의 생체신호들이 빠르게 생산되고 있다. 대표적으로 스마트 워치를 이용하여 가속도신호를 측정하고 이를 이용하여 개인의 운동량을 분석해주는 서비스[3]는 이미 대중화되어 있다. 하지만 이러한 서비스는 빠른 분석 결과가 필요하기 때문에 가속도 신호 raw-data를 이용하여 분석하지 않고, 분석에 필요한 특징 값들을 추출하여 분석에 이용한다. 따라서 가속도 신호를 측정 하지만 버려지는 데이터들이 많다. 비교적 측정하기 쉬운 가속도 신호 raw-data는 그 활용도가 매우 크다. 개인의 운동량 분석과 더불어 다양한 집단의 행동패턴 및 활동량 분석으로 특정 집단의 과로사, 특정 연령대의 과도한 운동에서 비롯되는 심근

* 정회원, 신라대학교 컴퓨터교육과

이 논문은 신라대학교 학술연구비 지원에 의하여 연구되었음.

접수일자 : 2017년 08월 29일

수정일자 : 2017년 09월 19일

게재확정일 : 2017년 09월 25일

교신저자 : 김태웅 e-mail : tmkim@silla.ac.kr

경색과 같은 원인 분석을 할 수 있는 하나의 지표로 사용될 수 있다. 따라서 손쉽게 측정할 수 있는 가속도신호의 raw-data를 서버로 전송하고 빅데이터 기반으로 다양한 분석을 하고 서비스할 수 있는 시스템이 필요하다.

본 논문에서는 일상생활에서 24시간 측정된 가속도 신호 raw-data를 서버로 전송하고, 이러한 데이터를 이용하여 집단 행동패턴 분석과 활동량 분석을 할 수 있는 시스템을 제안한다. 또한 집단의 시간, 요일, 연령, 성별 등의 기준으로 분석사례를 제시함으로써 효용성을 검증한다.

본 논문의 구성은 2장에서 관련연구로서 빅데이터, 생체신호 및 하둡플랫폼과 빅데이터 서비스 사례에 대해 기술하고, 3장에서는 제안한 분석 시스템구성 및 방법과 분석사례로 구성된다. 끝으로 4장에서는 제안한 시스템의 장점 및 향후연구과제로 결론을 맺는다.

II. 관련연구

1. 빅데이터

빅데이터(big data)는 기존의 데이터 단위를 넘어서는 방대한 양(Volume), 데이터의 생성과 흐름이 매우 빠르게 진행되는 속도(Velocity), 사진, 동영상 등 기존의 구조화된 데이터가 아닌 다양한(Variety) 형태의 정보 등 3가지 속성(3V)을 가지며 [4,5], 최근에는 가치(Value)까지 포함하여 네 가지 특징을 가진 데이터로 정의되는 것이 일반적이다. 세계 각국의 정부와 기업들은 빅데이터가 향후 국가와 기업의 성패를 가름할 새로운 경제적 가치의 원천이 될 것으로 기대하고 있으며, McKinsey[5], Gartner[6] 등은 빅데이터를 활용한 시장변동 예측과 신사업 발굴 등 경제적 가치창출 사례 및 효과를 제시하고 있다. 빅데이터란 큰 양의 데이터를 전제하며, '데이터'란 처음부터 특정 형식과 목적에 맞춰 작성된 구조화된 데이터만이 아니라, 비구조화 데이터인 SNS 메시지나 구글검색, 블로그, 이미지, 동영상 등이 분석과 활용의 대상으로 포함된다. 또한 넓은 의미의 빅데이터는 좁은 의미의 빅데이터 개념에 인적자원, 조직, 데이터처리·축적·분석기술(하둡, NoSQL, 기계학습, 통계분석 등)을 합친 개념이다. 빅데이터 분석기술은 통계, 데이터마이닝, 기계학습, 자연어처리, 패턴인식, 소셜 네트워크 분석, 비디오·오디오·이미지 프로세싱 등이 해당된다. 빅데이터의 활용, 분석, 처리 등을 포함하는 인프라에는 BI(Business Intelligence), DW(Data Warehouse), 클라우드 컴퓨팅, 분산 데이터베이스(NoSQL), 분산 병렬처리, 하둡(Hadoop) 분산파일시스템(HDFS), Map/Reduce 등이 있다.[7,8]

2. 생체신호

생체신호란 사람의 신체에서 발생하는 신호에 대하여 측정하는 순간의 값을 의미하며 전문적인 질병의 치료·관리가 이루어지는 의료 기관에서뿐만 아니라 최근에는 개인적으로 건강·질병 관리를 위한 정보로 활용 되고 있다. 일상생활 중 인체로부터 건강상태를 모니터링 할 수 있는 생체정보로는 체온, 심전도, 호흡, 폐음, 심음, 근전도, 피부저항, 목소리, 뇌파, 혈당, 체중, 혈압 등 매우 다양한 것들이 있다[9]. 또한 다양한 종류의 생체신호는 그 신호의 크기가 매우 작기 때문에 생체신호 측정 기술은 중요 계측 기술이 되고, 이를 계측하는 센서에는 산소포화도 센서, 심전도 센서, 혈압 센서, 가속 센서, 체온 센서, 호흡 센서 등이 있다[10].

최근 생체신호 측정 센서를 가진 생체신호의 측정이 가능한 웨어러블 디바이스의 등장으로 전문적인 지식 없이도 쉽게 생체신호를 활용한 건강 및 질병의 관리를 할 수 있게 되었다. 또한 병원이 아닌 일상생활에서 쉽게 측정이 가능한 웨어러블 디바이스들이 생산됨에 따라 인체로부터 발생하는 생체신호를 수집하는 도구가 많아졌으며 이것은 다양한 종류의 생체신호를 측정할 수 있음과 동시에 데이터를 빠르게 수집할 수 있다는 것을 의미한다.

웨어러블 디바이스를 통해 생성되는 생체신호의 생성속도(Velocity)는 매우 빠르며 생성되는 생체신호의 속도가 빠를수록 대용량(Volume)의 크기를 가지며 연속적으로 유입되는 가속도, 심전도, 호흡 등의 생체신호는 빅데이터의 특징인 비정형의 형태(Variety)가진다. 또한 생체신호를 통해 사용자에게 제공할 수 있는 수많은 가치(Value)를 가지므로 생체신호는 빅데이터로 볼 수 있으며 빅데이터 기반의 분석은 선택이 아닌 필수이다.

가속도 신호는 인체의 움직임을 측정할 수 있는 신호로 X축, Y축, Z축으로 구성되어 있으며 대부분의 웨어러블 디바이스에서 가속도신호를 이용하여 사람의 움직임을 추출한다. 스마트폰으로부터 추출할 수 있는 가속도 신호는 초당 50개의 신호를 저장하는데 24시간 측정 시에 17,280,000개의 데이터가 생성되므로 크기와 생성속도에서 빅데이터의 요소를 가지며 선형구조의 형태이다. 가속도 신호를 이용하여 얻을 수 있는 특징 값으로는 걸음 수, 소모 칼로리, 이동거리 등이 있으며 걷기, 뛰기, 자전거타기와 같은 행위와 앉기, 서있기와 같은 자세를 인지를 분석할 수 있다[11].

3. 하둡 플랫폼

하둡은 간단한 프로그래밍 모델을 사용하여 컴퓨터의 클러스터에 걸쳐 대규모 데이터의 분산처리를 허용하는 프레임워크

대[12]. 대용량 데이터의 분산저장과 다수의 서버 클러스터에서 일어나는 병렬처리를 하며, 파일을 나누어 분산 저장하는 하둡 분산파일시스템(HDFS)을 이용한다. 맵/리듀스(Map/Reduce)는 데이터 처리를 위한 프로그래밍 모델이며, 단순하지만 타고난 병행성을 내포하고 있다. 맵은 데이터를 가공 및 분류하는 역할을 수행하며, 리듀스는 전체 원소를 모아 하나의 결과를 출력하고자 할 때 사용한다. 이 두 함수의 구현을 통해 데이터 처리를 쉽게 병렬화할 수 있고, 범용적으로 활용 가능하다.

4. 빅데이터 서비스 사례

캐나다 온타리오 공과대병원에서는 IBM의 빅데이터 기술을 이용하여 미숙아 모니터링을 통한 감염 예방과 예측을 실시하고 있다. 이 미숙아 모니터링 시스템은 신생아의 혈압·체온·심전도와 혈중 산소 포화도 등 미숙아 모니터링 장비에서 생성되는 환자당 하루 9000만 건 이상의 생리학 데이터 스트림(data stream)을 실시간으로 분석하고 있다. 인큐베이터 안의 미숙아 및 신생아들의 신체 정보를 실시간으로 분석해 의료진이 신생아의 이상 징후를 알아낼 때보다 최소 6시간에서 24시간이나 먼저 감염 사실을 알아내고 위험 상황을 밝혀냄으로써 상태가 더 악화되기 전에 치료를 시작할 수 있다.

무엇보다 이러한 서비스가 가능하기 위해서는 한 종류의 데이터로 다양한 분석결과를 얻을 수 있도록 수집된 데이터가 raw-data 형태여야 한다. 특정한 하나의 분석결과를 얻기 위한 특징값들로 구성된 데이터는 다양한 분석결과를 얻을 수 없기 때문이다. 따라서 본 논문에서는 스마트폰을 통하여 가속도 신호를 24시간 수집하고 이를 서버로 전송하여 다양한 분석결과를 얻을 수 있는 시스템을 제안한다. 분석사례로는 모집단의 연령대별, 시간대별, 요일대별, 성별에 따른 활동량을 분석한다.

III. 빅데이터 기반 분석 시스템

이번 장에서는 본 논문에서 제안하는 시스템의 절차와 구조에 대해 설명한다. 제안하는 시스템은 생체신호의 전송, 하둡을 이용한 데이터의 저장, 데이터 변환 엔진, 데이터 분석 엔진, 서비스 프로바이더로 구성된다.

1. 분석 절차 및 시스템 구성

생체신호 분석 프로세스는 그림 1과 같이 다수의 사람들로 부터 측정된 생체신호 raw-data를 서버로 전송하고 서버는 빅데이터를 저장하는 하둡 플랫폼의 저장소인 HDFS(Hadoop Distributed File System)에 생체신호 raw-data를 저장한다. 저장된 생체신호 데이터 자체만으로는 의미가 없는 비정형의 형태를 가지고 있다. 따라서 이것을 특징값으로 변환한다. 이를

위하여 생체신호 분석알고리즘을 맵/리듀스 과정을 거쳐 특징값을 추출한다. 추출된 특징 값도 빅데이터에 해당하므로 하둡의 저장소인 HDFS에 저장한다. 끝으로 맵/리듀스를 통하여 결과를 SOAP기반으로 전송한다.

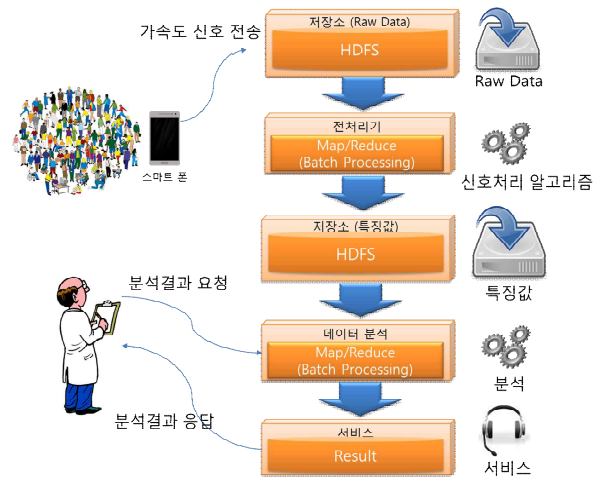


그림 1. 제안하는 시스템의 분석 절차 및 구성도

가. 전송 및 저장

생체신호의 전송은 다양한 종류의 스마트폰에서 전송을 할 수 있어야 하므로 표준 SOAP 메시지를 이용하여 전송한다. 서버로 전송된 생체신호는 빅데이터 저장소로 사용되는 하둡의 분산파일 시스템인 HDFS[11][13]에 저장한다. 그림 2는 이러한 절차를 나타내고 있다.

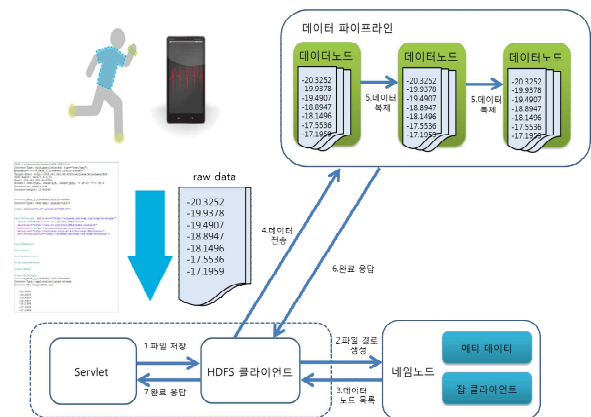


그림 2. 생체신호 전송 및 저장

나. 특징값 추출

가속도 신호와 같이 연속된 선형구조를 가지는 생체신호 raw-data는 생체신호를 생체신호 알고리즘을 거쳐 특징 값을

추출한다. 그림 3은 맵/리듀스에서의 생체데이터 분석 처리 과정을 나타내며 ①에서는 빅데이터 저장소인 HDFS에 저장된 생체신호 raw-data를 입력을 받는데 키는 데이터의 크기이고 값은 raw-data이다. ②에서 키와 값은 맵 메서드의 입력 데이터로 전달되며 ③의 맵과정에서는 생체신호 raw-data를 분석하는 알고리즘을 적용하여 의미를 가지는 특징값을 추출한다. 키 값으로는 생체신호 대상자의 기본정보를 삽입하고 생체신호 raw-data로부터 알고리즘을 적용하여 추출된 특징 값을 값으로 삽입하고 출력한다. ④는 맵과정을 거쳐 나온 출력 데이터를 정렬하며 ⑤에서 정렬된 출력 데이터를 리듀스 메서드의 입력 데이터로 전달한다. ⑥의 리듀스과정에서는 입력된 키에 따른 리스트로 전달되는 특징 값을 연산을 하여 출력한다. ⑦에서는 출력 데이터를 빅데이터 저장소인 HDFS에 저장한다.

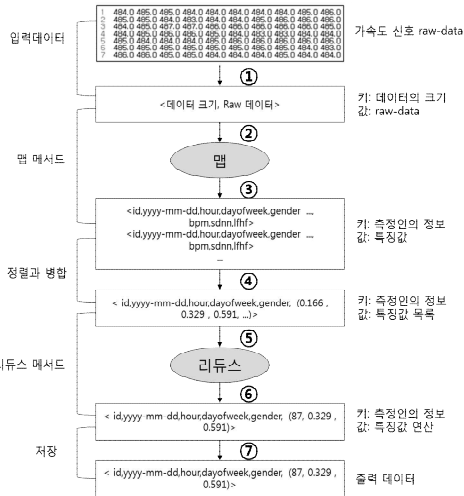


그림 3. 생체신호 추출 절차

다. 가속도 신호 분석

생체신호 raw-data로부터 특징 값을 추출하는 맵/리듀스 과정을 거친 생체신호 특징 값은 정형화되어 있으나 데이터베이스로 관리하기에는 부적합하다. 따라서 데이터 분석을 원하는 사용자에게 서비스를 제공하기 위해 그림 4와 같이 맵/리듀스 과정이 필요하다. 맵 과정에서 key-value쌍으로 생성하고, 리듀스 과정에서는 맵에서 출력된 값을 연산한다. 최종적으로 분석하고자하는 키값에 값을 매핑하여 출력데이터를 산출한다.

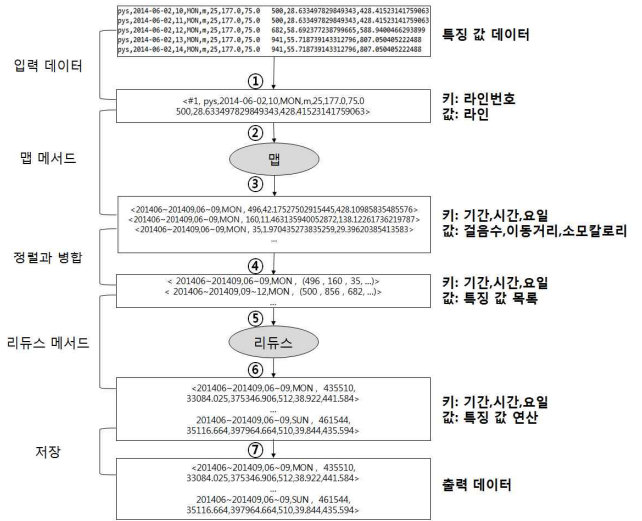


그림 4. 신호 분석을 위한 맵/리듀스 절차

2. 분석사례

스마트폰으로부터 추출된 가속도 신호를 빅데이터 기반으로 저장하고 특징값 추출 및 신호분석을 위한 맵/리듀스 과정을 거친 후 다음의 4가지에 대해 분석한다.

- a. 모집단의 시간대별 활동량 분포도
- b. 모집단의 연령대별 활동량 분포도
- c. 모집단의 요일대별 활동량 분포도
- d. 모집단의 성별에 따른 활동량 분포도

분석을 위하여 4개월간 24명, 연령은 10~50대의 연령분포를 가지며, 남녀 성별은 같게 하였다. 신호의 전송은 1시간 간격으로 서버에 자동전송하고, 1시간동안 측정된 raw-data의 용량은 8MB정도로 전체 300GB정도의 데이터를 샘플로 사용하였다.

가. 가속도 신호 특징값 추출

가속도 신호에는 x, y, z축 3축의 값을 가지고 있다. 3축의 값들을 이용하여 신체 움직임 감지할 수 있다. 하지만 측정하는 기기(스마트폰)의 위치(상의 주머니, 하의 주머니 등)에 따라서 3축의 신호들은 달라질 수 있으며 이러한 상황에 대해서도 알고리즘 분석이 가능해야 한다. 가속도 신호를 분석하기 위한 알고리즘은 zero crossing 알고리즘을 이용한다. 스마트폰으로부터 측정되고 하둡 플랫폼의 HDFS저장소에 저장된 가속도 신호의 raw-data는 그림 5와 같이 timestamp, x축, y축, z축으로 구성되며 웨이브형태의 의미가 없는 반정형의 수치 값이다.

1404194401200,8.171555,0.20385742,5.4515686
1404194401220,8.707733,0.6109009,5.133789
1404194401240,8.889832,1.1923065,4.542862
1404194401260,7.551468,1.3630981,5.792557
1404194401280,7.186676,1.3476257,7.7867126
1404194401300,5.6316986,3.4619904,8.161026
1404194401320,6.5731354,3.7779846,8.227081

그림 5. 가속도 신호의 형태(Timestamp, x축, y축, z축)

따라서 raw-data를 분석이 가능한 형태를 변환할 필요가 있으므로 맵/리듀스 과정을 통해 분석 가능한 특징 값 데이터로 정형의 데이터를 추출한다. 이러한 과정을 담당하는 맵은 알고리즘에 필요한 파라미터 요소인 성별, 키, 몸무게, 디바이스 종류를 추출하거나 id를 이용하여 기존의 데이터베이스로부터 추출한다. 그리고 이것을 특징값 추출 알고리즘을 통하여 얻어진 값들(걸음 수, 이동거리, 소모 칼로리)과 key-value 쌍으로 생성한다. 리듀스에서는 다수의 맵에서 분산으로 생성된 추출 데이터를 수집하여 처리한다. 그림 6과 같이 key 값에 설정된 id, 측정일, 측정요일, 성별, 나이, 키, 몸무게와 key를 가지는 value의 집합의 요소인 걸음 수, 이동거리, 소모 칼로리를 합하여 value값으로 출력한다.

```

hadoop@master:~/devel/hadoop-1.2.1/bin$ ./hadoop fs -cat /user/hadoop/result/Accres
ult_201309-201409/part-r-00000 | tail -10
pys,2014-09-01,14,MON,m,24,177.0,75.0 856,51.75749156766036,736.1065504745825
pys,2014-09-01,15,MON,m,24,177.0,75.0 876,52.06162318763836,740.4319776262034
pys,2014-09-01,16,MON,m,24,177.0,75.0 1280,71.55699830926008,1017.6995371980393
pys,2014-09-01,17,MON,m,24,177.0,75.0 2408,143.16620912516467,2036.1416537574678
pys,2014-09-01,18,MON,m,24,177.0,75.0 3424,207.1866037905743,2946.6539399824887
    
```

그림 6. 특징 값 추출 맵/리듀스로 추출된 정형화된 가속도신호

나. 분석결과

```

<?xml:namespace="http://schemas.xmlsoap.org/soap/envelope/">
<?xml:namespace="http://service.wellness/">
<getFusionDataInfoResponse>
  <return>
    <avg_calorie>33.5141119165747</avg_calorie>
    <avg_distance>499.377302609463</avg_distance>
    <avg_steps>544</avg_steps>
    <factor>0.9</factor>
    <feature>time</feature>
    <term>201406-201409</term>
    <total_calorie>7624.46046112323</total_calorie>
    <total_distance>1067833.8812323</total_distance>
    <total_steps>1237600</total_steps>
  </return>
  <return>
    <avg_calorie>146.78358053341</avg_calorie>
    <avg_distance>1446.70105123233</avg_distance>
    <avg_steps>1676</avg_steps>
    <factor>1.8</factor>
    <feature>time</feature>
    <term>201406-201409</term>
    <total_calorie>333932.66546524</total_calorie>
    <total_distance>3291244.89231567</total_distance>
    <total_steps>3812900</total_steps>
  </return>
</getFusionDataInfoResponse>
</Body>
</Envelope>
    
```

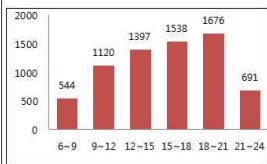
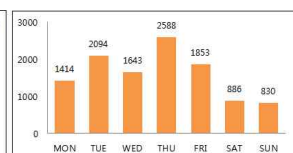


그림 7. 시간대별 활동량 분석 응답 메시지 및 그래프

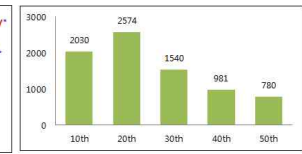
```

<?xml:namespace="http://schemas.xmlsoap.org/soap/envelope/">
<?xml:namespace="http://service.wellness/">
<getFusionDataInfoResponse>
  <return>
    <avg_calorie>2094</avg_calorie>
    <avg_distance>1643</avg_distance>
    <avg_steps>1414</avg_steps>
    <factor>1.8</factor>
    <feature>time</feature>
    <term>201406-201409</term>
    <total_calorie>3291244.89231567</total_calorie>
    <total_distance>3291244.89231567</total_distance>
    <total_steps>3812900</total_steps>
  </return>
</getFusionDataInfoResponse>
</Body>
</Envelope>
    
```



```

<?xml:namespace="http://schemas.xmlsoap.org/soap/envelope/">
<?xml:namespace="http://service.wellness/">
<getFusionDataInfoResponse>
  <return>
    <avg_calorie>2030</avg_calorie>
    <avg_distance>2574</avg_distance>
    <avg_steps>1540</avg_steps>
    <factor>1.8</factor>
    <feature>time</feature>
    <term>201406-201409</term>
    <total_calorie>7624.46046112323</total_calorie>
    <total_distance>1067833.8812323</total_distance>
    <total_steps>1237600</total_steps>
  </return>
</getFusionDataInfoResponse>
</Body>
</Envelope>
    
```



```

<?xml:namespace="http://schemas.xmlsoap.org/soap/envelope/">
<?xml:namespace="http://service.wellness/">
<getFusionDataInfoResponse>
  <return>
    <avg_calorie>2245</avg_calorie>
    <avg_distance>1503</avg_distance>
    <avg_steps>1503</avg_steps>
    <factor>1.8</factor>
    <feature>time</feature>
    <term>201406-201409</term>
    <total_calorie>7624.46046112323</total_calorie>
    <total_distance>1067833.8812323</total_distance>
    <total_steps>1237600</total_steps>
  </return>
</getFusionDataInfoResponse>
</Body>
</Envelope>
    
```

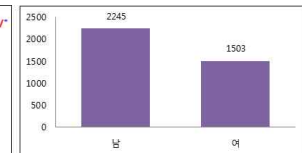


그림 8. 요일, 연령별, 성별에 따른 활동량 요청 메시지와 응답 결과 그래프

그림 7은 주된 움직임 시간인 오전 6시부터 밤 12시까지 동안 3시간 단위로 측정사람들의 평균 걸음수, 소모 칼로리, 이동거리, 총 걸음수, 칼로리 소모량, 이동거리를 추출하였으며, 이를 통해 집단의 사람들이 어떤 시간에 가장 많은 활동량을 가지는지 유추할 수 있다. 그림 8은 각각 오후 6시에서 9시까지의 요일별, 연령별, 성별에 따른 활동량을 분석한 결과이다. 이것으로 목요일의 활동량과 어느 연령보다 20대, 그리고 여성에 비해 남성의 활동량이 많다는 것을 알 수 있다.

IV. 결론

본 논문에서 우리는 가속도 신호 raw-data를 이용하여 빅데이터 기반의 분석 시스템을 제안하고 4개월간 24명으로부터 수집된 데이터를 이용하여 분석 사례를 들었다. 특징기간 동안 집단의 시간을 기준으로 한 활동량 분석으로 3시간 단위로 어떤 시간대에 가장 활동량이 많은지, 집단의 연령별로 어느 연령대가 가장 활동량이 많은지, 집단의 요일별, 집단의 성별 그리고 어떤 특정시간에 연령, 요일 그리고 성별로 가장 활동량이 많은지를 분석하고 서비스 해 보았다. 클라이언트가 수집한 데이터를 SOAP메시지 형태로 서버에 전송하면 서버는 HDFS에 저장과 동시에 분석에 필요한 특징값을 추출하고, 추출된 특징값을 다시 HDFS에 저장한다. 클라이언트가 SOAP 메시지를 이용하여 특정한 분석을 요청하면 서버는 맵/리듀스를 이용하여 분석한 결과를 다시 SOAP메시지 형태로 응답한다.

가속도신호에 대한 빅데이터 서비스는 좁은 범위에서는 한 집단의 사람들의 평소 움직임상태에서 넓은 의미로는 국가 전체의 사람들의 활동량에 대한 데이터로 집단의 사람들의 움직임을 분석하고 사람들의 행동패턴을 분석하는 기반연구가 될 수 있다. 또한 다양한 생체신호를 분석하고 서비스 할 수 있으며, 측정단말기와 서비스를 요구하는 단말기의 종류에 독립적인 플랫폼을 구현함으로써 상호운용성이 확보된 플랫폼이라는 장점을 가진다.

현재 심전도나 호흡신호를 이용한 분석과 서비스에 대한 연

구를 진행 중에 있으며, 이것은 질병예측과 같은 의료분야에서 빅데이터 기반의 생체신호 분석 플랫폼의 유용한 선행연구가 될 것이다.

향후 지속적인 데이터의 수집을 통하여 보다 큰 모집단에 대한 데이터와 보다 다양한 활동성 분석지표를 설정하여 직업군에 대한 과로사 원인을 분석하는 연구를 진행할 것이다.

REFERENCES

- [1] 김성우, 김각규, 윤봉규, “국방분야 빅데이터 분석의 활용가능성에 대한 고찰,” *한국경영학회*, 제29권, 제2호, 1-19쪽, 2014년 6월
- [2] 신신애 “공공기관 빅데이터 활용 현황 및 과제,” *한국정보화진흥원*, 제57권, 제5호, 398-404쪽, 2014년
- [3] 이영훈, 김용일, “Mi Band와 MongoDB를 사용한 생체정보 빅데이터 시스템의 설계,” *스마트미디어저널*, 제5권, 제4호, 124-130쪽, 2016년 12월
- [4] Magoulas, Roger, B. Lorica, “Big data: Technologies and techniques for large scale data,” Release 2.0, 2009.
- [5] McKinsey, “Big data, The next frontier for innovation, competition, and productivity,” McKinsey Global Institute, 2011.
- [6] Gartner, Cool Vendors in Wearable Electronics for Health and Fitness, 2013.
- [7] W. Peter, Big Data Glossary, O'Reilly Media, 2011.
- [8] 장상현, “빅데이터와 스마트교육,” *한국정보과학회지*, 제30권, 제6호, 59-64쪽, 2012년 6월
- [9] S.H. Kim, “Biological signal monitoring technique for ubiquitous healthcare,” *IT SoC Mag*, vol. 25, pp. 40-47, 2008.
- [10] 김준수, 이가영, 유철승, 김태웅, 이상훈 김희철, “의복형 생체신호 모니터링을 이용한 건강증진 시스템,” *Korean Journal of Health Promotion*, 제11권, 제1호, 1-8쪽, 2011년
- [11] Y. Jun, “Toward physical activity diary: motion recognition using simple acceleration features with mobile phones,” Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics, *ACM*, 2009.
- [12] Apache™ Hadoop®, <http://hadoop.apache.org/>
- [13] Tom White, Hadoop The Definitive Guide 2nd Edition, *O'Reilly Media*, pp. 41-47, 2009.

저자 소개



김태웅(정회원)

1994년 인제대학교 전산학과 학사 졸업.
1998년 인제대학교 전산학과 석사 졸업.
2009년 인제대학교 전산학과 박사 졸업.

2007년~2009년 인제대학교 컴퓨터공학과 연구교수
2014년~현재 신라대학교 컴퓨터교육과 교수

<주관심분야 : 소프트웨어공학, 역공학, 리팩토링, 헬스케어, 빅데이터>