

R 을 이용한 수면 데이터 분석 및 시각화

장유희*, 유치현**, 임효상*

*연세대학교 원주캠퍼스 컴퓨터정보통신공학부

**피에치엠컴퍼니(주)

e-mail : yuheejang@yonsei.ac.kr, chihun@phmcompany.com, hyosang@yonsei.ac.kr

Analysis and Visualization of Sleep Data by Using R

Yuhee Jang*, Chi-Hun Yoo**, Hyo-Sang Lim*

*Computer and Telecommunications Engineering Division, Yonsei University

**PHM & Company Inc.

요 약

최근 인구 고령화 시대에 맞춰 개인건강에 대한 관심이 높아지고 있다. 이에 따라 헬스케어에 대한 관심이 집중되고 있으며, 스마트 기기나 웨어러블(wearable) 기기를 이용한 다양한 헬스케어 서비스들이 등장하였다. 특히 건강한 삶을 유지하는데 필수적으로 요구되는 수면에 관한 서비스들이 존재하지만, 사용자에게 실질적인 효용성을 제공하는 것에는 한계를 드러낸다. 본 논문에서는 사용자 자신의 수면 상태를 점검하여 올바른 수면 습관을 유도할 수 있도록 사용자의 수면주기를 시각화시키는 연구를 수행한다. 이를 위해 데이터 통계와 시각화에 적합한 언어인 R 을 사용하여 사용자의 수면 데이터를 분석하여 시각화한다.

1. 서론

최근 개인건강에 대한 사람들의 관심이 높아지면서 헬스케어 부문에 대한 수요가 증가하고 있다. 또한 헬스케어에 IT 기술을 접목시켜 사용자의 삶의 질을 향상시키기 위한 다양한 서비스들이 등장하였다. 특히 스마트폰은 사용자에게 휴대성, 이동성을 제공할 뿐만 아니라 GPS, 가속도 센서, Bluetooth, 심장 박동 센서 등을 비롯한 다양한 센서들을 탑재하고 있어 헬스케어 서비스에서 중요한 역할을 한다[1]. 스마트폰에 연결된 웨어러블(wearable) 기기에서 측정된 데이터를 스마트폰의 앱을 통해 사용자에게 제공하거나, 스마트폰의 센서를 통해 직접 사용자의 신체활동을 감지하여 서비스를 제공한다.

스마트폰에 탑재된 센서를 통해 수집된 사용자의 신체활동 정보를 활용하는 다양한 헬스케어 서비스가 있다. 대표적인 예로 만보 계를 이용해 걸음걸이 수를 측정하여 하루 동안 사용자가 소모한 칼로리 양을 알려주는 서비스가 있다. 또한 사용자가 수면 시 뒤척이는 정도를 가속도 센서를 통해 측정하여 사용자의 수면 상태를 알려주는 서비스가 있다. 이런 서비스는 사용자의 수면주기를 파악해 사용자가 잠에서 깨었을 때 가장 상쾌함을 느낄 수 있는 시간에 사용자를 깨워 수면의 효율성을 높이고자 한다.

하지만 대부분의 수면 상태를 알려주는 서비스의 경우 사용자에게 실질적인 효용성을 제공하는 것에

한계를 드러낸다. 스마트폰, 웨어러블 기기 등으로부터 수면 시 측정된 사용자의 뒤척임 정도를 빠르고 정확하게 측정하는 것은 가능하지만 단순히 측정 결과를 대쉬보드나 그래프 형태로 제공한다. 따라서 측정 결과가 의미하는 바가 무엇인지, 그 결과물을 통해 사용자가 어떤 이득을 얻을 수 있는지는 명확하게 제공하지 못하고 있다.

본 논문에서는 사용자의 수면 시 뒤척임 정도를 나타내는 수면 데이터를 스마트폰의 가속도 센서를 이용해 수집하고, 수집된 수면 데이터를 분석하여 사용자의 수면주기를 파악하는 연구를 진행한다. 실제 사용자의 정확한 수면주기를 측정하기 위해서는 고가의 의료 장비를 이용해 측정된 근육의 긴장도, 안구운동, 뇌파 등의 다양한 생체 신호 정보가 필요하다. 수면 중 뒤척임 움직임 정보만을 통해 사용자의 정확한 수면주기를 파악하는 것은 어려움이 따른다. 따라서 본 논문에서 수행하는 분석은 수면 중 측정된 움직임을 통해 사용자의 수면주기를 비교적 정확하게 파악하여 시각화 시키기 위한 선행 작업이다. 데이터 통계와 그래픽에 최적화된 언어인 R[2]을 통해 분석을 수행하고, 각 과정의 분석 결과를 시각화 시키는 것에 초점을 맞춘다.

논문의 구성은 다음과 같다. 2 장에서는 본 연구에서 다루는 관련 내용에 대해 소개하고, 3 장에서는 수면 데이터의 분석 결과 및 시각화에 대해 기술한다. 마지막으로 4 장에서는 결론 및 향후 연구로 논문을 끝맺는다.

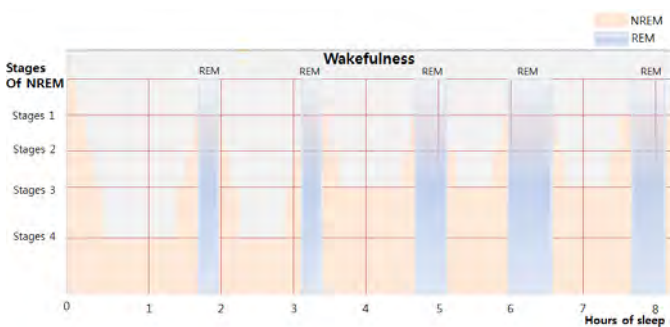
본 논문은 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 "바이오제조 GMP 기술인력양성사업"의 지원을 받아 수행하였습니다 (N0000961).

2. 관련 내용 소개

본 논문에서는 스마트폰의 가속도 센서를 통해 수면 중에 수집된 사용자의 뒤척인 정도를 측정하고, R을 활용해 수면주기를 시각화시키는 연구를 진행한다. 본 연구에서 수행한 분석 과정을 설명하기에 앞서 관련된 내용을 기술한다.

2.1 수면주기

인간의 수면은 생리적으로 한가지 상태로 유지되지 않으며, REM(Rapid Eye Movement)수면과 NREM(Non-REM)수면의 두 가지 상태로 구분할 수 있다[3]. NREM 수면은 깊이에 따라 그림 1 과 같이 4 단계로 나눌 수 있으며 높은 단계일수록 더 깊은 수면 상태를 나타낸다. 또한 높은 단계일수록 각성 상태로의 전환을 위해 더 강한 자극을 필요로 한다.



(그림 1) 수면주기

수면주기는 NREM 수면과 REM 수면이 교대로 나타나게 되며, 개인차가 존재하지만 수면장애 없이 생활하고 있는 정상인의 한 수면주기는 약 90~120 분으로 알려져 있다[4]. 그림 1 에서처럼 사람이 잠에 들면 NREM 수면(약 90 분)으로 시작해 REM 수면(약 30 분)을 정점으로 한 주기를 이루게 된다.

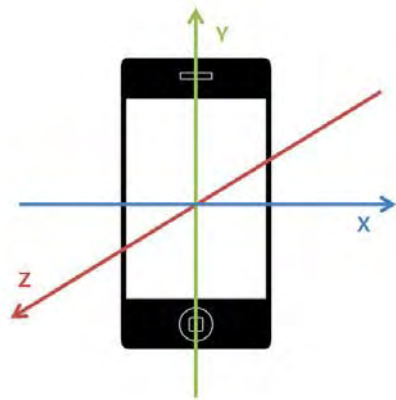
REM 수면은 뇌는 깨어있지만 몸은 쉬고 있는 상태를 의미하며, 각성 시 외부 자극에 대해 머리와 몸이 빠르게 반응 한다. NREM 수면은 신체는 깨어있지만 뇌는 쉬고 있는 상태를 말하며, 각성 시 뇌는 알게 잠들어 있지만 육체는 깊이 잠들어 있어 균형이 잡히지 않은 상태이기 때문에 기분이 상쾌하지 못한 느낌을 받게 된다. 따라서 상쾌하게 잠을 깨기 위해서는 수면의 한 주기가 끝난 시점에서 눈을 뜨는 것이 가장 좋다.

이런 수면주기의 각 상태에 따른 생리적 특성을 이용해 수면의 질을 높이 위한 헬스케어 앱들이 존재한다. 대표적인 예로 하루 단위로 사용자의 수면 상태를 파악해 수면의 품질을 알려주는 앱인 “SleepCycle”이 있다[5]. 또 다른 예로는 사용자의 수면주기를 파악하여 일어나기 가장 좋은 시간에 사용자를 깨워주는 스마트 알람 앱인 “SleepBetter”가 있다[6]. 이처럼 사용자에게 보다 나은 수면의 질을 제공하기 위해서는 사용자의 수면주기를 정확하게 판별해 내는 것이 선행적으로 이루어져야 한다.

2.2 가속도 센서

사람의 수면주기를 파악하기 위해서 수면 중 변화하는 다양한 생리적 기능 및 신호가 이용된다. 대표적으로 이용되는 생리적 기능 및 신호로는 근육의 긴장도, 안구운동, 뇌파, 호흡, 맥박, 신체의 움직임 등이 있다[7,8]. 하지만 이런 정보를 측정하기 위해서는 대부분 고가의 장비들이 필요하고, 일반 사용자들의 수면주기를 파악하기 위해 사용하기에는 부적합하다. 따라서 본 논문에서는 스마트폰의 가속도 센서를 통해 사용자가 수면 시에 뒤척인 정도를 측정하고, 측정된 정보를 바탕으로 수면주기를 파악한다.

가속도 센서는 스마트폰의 단위시간 동안에 움직임을 측정하는 센서이다. 그림 2 는 스마트폰의 x, y, z 축을 나타낸다. 각 축은 스마트폰의 기울기와 가속도 두 가지 값을 측정하여 하나의 벡터 값으로 출력해주는 역할하며, 스마트폰의 급격한 동작 변화나 받는 충격에 따라 값이 지속적으로 변한다. 따라서 사용자가 수면을 취하는 자리 근방에 스마트폰을 배치하면 수면 시 사용자의 뒤척인 정도를 효과적으로 감지할 수 있다.



(그림 2) 스마트폰의 3축 가속도 센서

2.3 R

무료 언어인 R 은 데이터 통계 분석과 분석 결과를 시각화하기 위해 최적화된 언어이다. 데이터 통계 분석이란 축적된 데이터에서 통계적 분석 방법을 통해 특징을 찾아내는 것을 말한다. 대부분의 분석작업은 반복적으로 수행되며, 수행된 결과를 즉각적으로 확인할 수 있어야 한다. 시각화란 데이터를 사람이 직관적으로 쉽게 이할 수 있도록 하는 가장 효과적인 방법 중 하나이며, 분석 결과를 시각적으로 표현하여 전달하는 과정을 말한다.

R 은 데이터 통계 분석과 시각화에 대해 강점을 갖는 다양한 특성들을 가지고 있다. 먼저, 데이터 통계 분석을 위한 R 의 특성을 설명한다. R 은 함수형 언어의 특성을 갖기 때문에 간결하게 코드를 구현할 수 있고, 객체 지향적 요소들을 포함하고 있어 코드의 재사용 및 모듈화에 용이하여 반복적인 분석작업을 큰 노력 없이 수행할 수 있다. 또한 대화식(interactive)으로 수행되어 분석 결과를 즉시 확인하는 것이 가능하다. 다음으로 R 은 시각화를 위해 강력한 기능을 제

공하는 다양한 패키지를 CRAN(Comprehensive R Archive Network)라는 패키지 배포 시스템을 통해 무료로 제공한다. 또한 패키지를 로드(load)하는 간단한 명령어만 입력하면 추가적인 설정 없이 패키지를 즉시 사용하는 것이 가능하다. R은 앞서 언급한 다양한 특성들로 인해 빅 데이터 분석 플랫폼에서 데이터 분석 툴로 사용되고 있으며 주로 분석한 결과를 하나의 정보로 시각화하여 전달하는 역할을 한다. 이를 통해 분석 결과를 직관적으로 쉽게 이해하는 것이 가능하다.

본 논문에서는 사용자의 수면주기를 파악하기 하루 단위로 수집된 실험참가자의 수면 데이터를 반복적으로 분석하였다. 이를 위해 R을 이용해 수면 데이터를 다양하게 분석해 보았으며, 분석의 각 과정에서 생성되는 중간 결과로부터 의미를 직관적으로 발견하기 위해 분석결과를 시각화 시켰다.

3. 분석 과정

사용자의 수면 데이터를 분석하는 과정은 3 단계로 이루어진다. Step1에서는 스마트폰의 가속도 센서를 이용해 사용자가 수면 중 뒤척인 정도를 측정하는 수면 데이터와 수면 데이터 수집 방법에 대해 설명한다. Step2에서는 수집된 수면 데이터로부터 사용자의 수면주기를 파악하기 위해 수면 데이터 집계 및 분석을 수행한다. 마지막으로 Step3에서는 분석 결과를 토대로 실험 대상자의 수면주기를 시각화 시킨다.

3.1 Step1: 수면 데이터 수집

실험 참가자의 수면주기를 파악하기 위해 두 개의 필드로 구성된 수면 데이터를 수집한다. 첫 번째 필드는 수면 데이터의 수집 시점을 의미하는 타임스탬프(timestamp)이며, 두 번째 필드는 수면 중 사용자의 뒤척인 정도를 나타내는 충격 값이다. 이러한 수면 데이터를 수집하기 위해 안드로이드 스마트폰 앱을 구현하였고, 실험 참가자는 수면 직전에 앱을 실행한 뒤 침대 머리 근처에 놓고 수면을 취하게 된다. 활성화된 앱은 실험 참가자가 잠든 동안 뒤척이는 움직임을 일정시간 간격으로 가속도 센서를 통해 수집한다. 그리고, 실험 참가자가 잠에서 일어난 후 앱을 종료하게 되면 잠든 동안 수집된 충격 값을 csv 형식으로 출력하게 된다. 수면 데이터는 약 3주에 걸쳐 수집하였으며, 충격 값은 스마트폰 가속도 센서의 각 축의 벡터 값을 식(1)과 같이 계산하였다.

$$\text{충격 값} = \sqrt{x^2 + y^2 + z^2} \quad \text{식(1)}$$

그림 3은 구현한 앱을 통해 수집한 실험 참가자의 수면 데이터의 일부를 나타낸다. 실험 참가자의 하루 평균 수면 시간은 약 5시간이다. 수집 간격을 20ms로 설정하였기 때문에 평균적으로 하루 마다 수집되는 수면 데이터는 약 100만개 이상의 레코드로 구성

된다. 분석을 수행하기에는 데이터 양이 너무 크기 때문에 하루 마다 측정된 수면 데이터에 대해 집계를 수행한다.

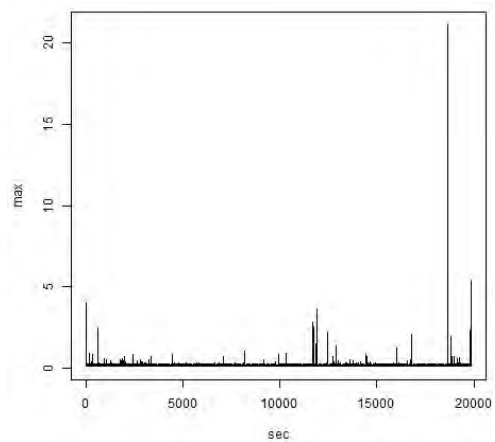
타임스탬프	충격 값
...	
10420617	1.182736
10420637	0.715866
10420658	1.142034
10420679	2.157176
...	

(그림 3) 수면 데이터 예시

3.2 Step2: 수면 데이터 집계 및 분석

수면 데이터 집계는 20ms 마다 수집된 수면 데이터를 일정한 t_1 시간 구간 단위로 묶는 과정이다. t_1 은 시스템 설정 값으로 수십 ms에서 수천 ms가 된다. 수면 데이터의 전체적인 분포나 동향을 파악하기 위해서는 각 시간 구간에서 뒤척인 정도를 나타낼 만한 통계 정보가 필요하다. 따라서 충격 값의 합, 평균 값, 중앙 값 등 다양한 요약 통계 정보를 산출해 보았지만 본 분석 과정에서는 충격 값의 최대값을 이용해 분석한 과정을 기술한다.

그림 4는 하루에 수집된 수면 데이터에 대해 t_1 시간 구간 단위로 집계를 수행한 결과를 R을 사용하여 시각화 시킨 것이다. 시각화를 통해 t_1 시간 구간 단위에서 충격 값의 최대 값 분포를 직관적으로 파악할 수 있으며, 특정 구간에서 최대 값이 비정상적으로 크게 나타난 것을 확인할 수 있다. 비정상적으로 큰 최대 값은 실험 참가자의 정확한 수면주기를 파악하는데 불필요하다. 따라서 상한 값을 지정하여 상한 값보다 크면서 가장 작은 값을 측정하고, 비정상적으로 크게 발생한 최대 값을 상한 값으로 변경시킨다.

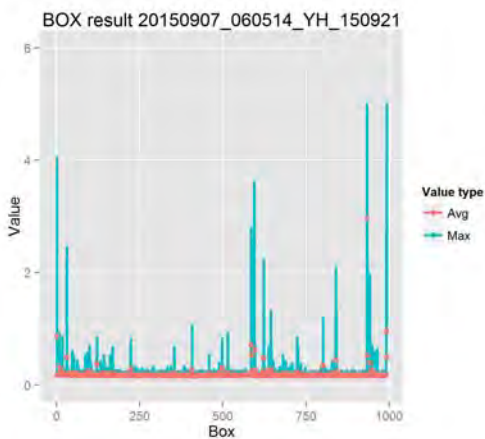


(그림 4) t_1 단위 구간 집계 수행 결과 시각화 예

3.3 Step3: 분석 결과 시각화

그림 5는 R을 사용하여 Step2에서 수행한 결과를

t_2 시간 구간 단위로 2 차 집계하여 시각화시킨 실험 참가자의 수면주기 예시이다. t_2 는 t_1 보다 큰 시간 구간 단위로 수 초에서 수분 사이의 값이다. 실선으로 표현된 값은 t_2 단위 구간 내의 최대 충격 값을 의미하며, 점으로 표현된 값은 t_2 단위 구간 내의 평균 충격 값을 나타낸다. 각 단위 구간 별 최대 충격 값과 평균 충격 값의 경우 차이가 소수점 단위로 굉장히 작기 때문에 통계 값을 통해 바로 이해하기에는 어려움이 있지만 R 을 사용하여 시각화시킨 결과를 통해 각각의 결과 값이 어떤 차이가 있는지 직관적으로 이해하는 것이 가능하다.



(그림 5) 최대 충격 값을 이용한 수면주기 예시

충격 값의 최대값을 이용한 분석 과정에서 하루 단위로 수집된 실험 참가자의 수면주기를 시각화 시켜 파악해본 결과 수면 시 스마트폰의 위치와 수면 환경 (침대, 바닥 등)에 따라 수집되는 수면 데이터의 충격 값의 최대값 분포가 달라지는 양상을 관찰하였다. 따라서 좀 더 정확한 실험 참가자의 수면주기를 파악하기 위해서는 충격 값의 분포에 따라 상계 값을 조정하는 과정이 추가적으로 요구된다는 직관을 얻을 수 있다.

약 3 주에 걸쳐 수집된 실험 참가자의 총 수면 데이터는 약 2,000 만개 이상으로, 일일이 수면 데이터를 살펴보는 것은 불가능에 가깝다. 이를 위해 R 에서 제공하는 요약 통계 기능을 이용해 다양한 통계정보를 산출해 보았고, 시각화 기능을 이용해 각 분석 과정의 결과를 시각화 하였다. R 을 분석 및 시각화에 사용함으로써 분석에 요구되는 다양한 요약 통계 정보를 효율적으로 산출할 수 있으며, 이해하기 어려운 숫자의 나열인 경우가 대부분인 요약 통계 정보를 시각화 함으로써 보다 직관적이고 포괄적으로 분석 결과를 이해할 수 있다.

4. 결론 및 향후 연구

본 논문에서는 스마트폰의 가속도 센서를 활용해 사용자의 수면 데이터를 수집하였다. 또한 데이터 통계 및 데이터 시각화에 최적화된 언어인 R 을 이용해 수면 데이터를 분석하였고, 분석한 결과를 바탕으로

사용자의 수면주기를 시각화 시켰다. 향후 연구로는 보다 정확한 수면 주기를 파악하기 위해 다양한 요약 통계 정보를 종합적으로 활용할 것이며, 사용자 별로 수집된 수면 데이터를 바탕으로 사용자 개인의 수면 주기를 모델링하여 사용자 별 수면의 품질을 측정하는 연구를 진행할 계획이다.

참고문헌

- [1] Siewiorek, Dan. "Generation smartphone." *Spectrum*, IEEE 49.9 (2012): 54-58.
- [2] Ihaka, Ross, and Robert Gentleman. "R: a language for data analysis and graphics." *Journal of computational and graphical statistics* 5.3 (1996): 299-314.
- [3] Carskadon MA, Dement WC. Normal human sleep: An overview. In Kryger MH, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*. 5th ed. St. Louis: Elsevier Health Sciences; 2011:16-26.
- [4] Chae, Kyu Young. "Physiology of sleep." *Korean Journal of Pediatrics* 50.8 (2007): 711-717.
- [5] SleepCycle, <http://www.sleepcycle.com>.
- [6] SleepBetter, <https://www.runtastic.com/en/apps/sleepbetter>.
- [7] Sinton, Christopher M., and Robert W. McCarley. "Neurophysiological mechanisms of sleep and wakefulness: a question of balance." *Seminars in neurology*. Vol. 24. No. 3. 2004.
- [8] Koo, Dae Lim, and Juhan Kim. "The physiology of normal sleep." *Hanyang Medical Reviews* 33.4 (2013): 190-196.