

PHR 기반 개인 맞춤형 건강정보 탐사 알고리즘 설계

신문선

건국대학교 과학기술대학 컴퓨터공학과

Design of knowledge search algorithm for PHR based personalized health information system

Moon-Sun SHIN

Dept. of Computer Engineering, Konkuk University

요약 PHR(Personal Health Record)기반 헬스케어 서비스 플랫폼 지능화를 위해서는 사용자 맞춤형 건강정보 제공 서비스가 필요하다. 본 논문에서는 개인 맞춤형 건강정보 추천을 위해서 온톨로지 기반 건강 정보 모델을 제안하였다. 또한 기계학습과 데이터마이닝 기법을 적용한 유사 건강정보 탐사 알고리즘을 설계하였다. 기존의 데이터마이닝 기법중 연관규칙 알고리즘을 확장하여 속성을 기반으로 연관규칙 탐사를 수행하여 지식탐사의 연관성을 높이고 효율적인 탐사시간을 제공할 수 있도록 하였다. 머신러닝의 한 기법인 K근접이웃 알고리즘을 적용하여 사용자 프로파일별 그룹화를 수행하고 유사 패턴의 사용자 프로파일을 검색할 수 있도록 하였다. 이는 사용자의 질환과 건강상태에 따른 맞춤형 건강정보 탐사 수행의 효율성을 높인다. 제안된 알고리즘은 개인 맞춤형 헬스케어 서비스 플랫폼에서 추론과정에 적용되어 사용자에게 개인맞춤형건강정보를 추천하는 것을 가능하게 한다. 이는 고령화사회에서 스마트한 자가 건강관리에 활용될 수 있다.

주제어 : 개인건강기록, 온톨로지, 기계학습, 속성기반 연관규칙, PHR기반 건강정보서비스시스템

Abstract It is needed to support intelligent customized health information service for user convenience in PHR based Personal Health Care Service Platform. In this paper, we specify an ontology-based health data model for Personal Health Care Service Platform. We also design a knowledge search algorithm that can be used to figure out similar health record by applying machine learning and data mining techniques. Axis-based mining algorithm, which we proposed, can be performed based on axis-attributes in order to improve relevance of knowledge exploration and to provide efficient search time by reducing the size of candidate item set. And K-Nearest Neighbor algorithm is used to perform to do grouping users by according to the similarity of the user profile. These algorithms improves the efficiency of customized information exploration according to the user's disease and health condition. It can be useful to apply the proposed algorithm to a process of inference in the Personal Health Care Service Platform and makes it possible to recommend customized health information to the user. It is useful for people to manage smart health care in aging society.

Key Words : Personal Health Record, Ontology, Machine Learning, Axis based association rule, PHR based Health Information Service System

* This paper was supported by Konkuk University.

Received 25 February 2017, Revised 27 March 2017

Accepted 20 April 2017, Published 28 April 2017

Corresponding Author: Moon-Sun Shin(Konkuk University)

Email: msshin@kku.ac.kr

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

최근 ICT 융합기술과 사물인터넷 기술을 활용한 헬스케어 서비스는 질병의 치료는 물론, 진단, 관리적 측면의 산업을 활성화 시키는 한편 전통적 치료 산업의 비중을 감소시켜 질병치료의 시장동향의 변화를 가져오고 있다. 이는 개인 환자의 치료비 감소로 국가 전체적인 의료지원비용 감소 효과가 있으며 진료의 질은 높아지는 경제적 파급효과가 있다[1,2,3].

<Table 1> Prospect of Health Care Industry (2012 Samsung Economy Center)

(unit: billion dollar,%)

| category | prophylazy | diagnosis | medical treatment | Post-disease management | total |
|-----------------------|------------|------------|-------------------|-------------------------|-------------|
| 2010 | 2,140(5) | 5,700(16) | 24,240(58) | 3,580(10) | 35,640(100) |
| 2015 | 3,980(6) | 9,190(19) | 31,420(65) | 5,100(10) | 49,690(100) |
| 2020 | 6,880(10) | 14,400(21) | 39,110(57) | 8,230(12) | 68,600(100) |
| growth rate 2010*2020 | 12.4 | 9.7 | 4.9 | 8.7 | 6.8 |

<Table 2> Health Care Service Trend (2015 Government BT Policy)

| category | Tel-Health 1900*1999 | e-Health 2000*2005 | u-Health 2000*2010 | smart-Health 2010* |
|----------|-------------------------|------------------------|--------------------------------|--|
| provider | hospital | hospital | hospital, ICT industry | hospital, ICT industry, service provider |
| user | Medical Staff | Medical Staff, Patient | Medical Staff, Patient, People | Medical Staff, Patient, People |
| system | HIS, PACS | EMR System | EHR Monitoring | PHR based Service |

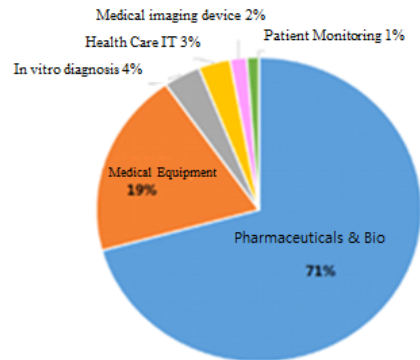
의료서비스의 환경변화는 의료기관의 치료가 중심인 전통적 의료서비스는 IT·BT 기술의 발전에 따라 이용자와 예방·관리 목적의 4P(주도형 헬스케어 서비스로 진화하였으며 최근 스마트화 시대의 도래와 함께 이용자(환자, 일반인 등) 개인별 맞춤형으로 지능화된 스마트 헬스케어가 각광을 받고 있다[4, 5]. <Table 1>은 삼성경제 연구소에서 발표한 헬스케어 산업의 전망에 대한 데이터이며 <Table 2>는 2015년 정부 BT 정책관련 자리에서 분석된 헬스케어서비스 발전 전망에 대한 데이터로 스마트헬스 및 예방 사후 관리목적의 헬스케어서비스로 진화하고 있음을 시사하고 있다[1].

[Fig. 1] 과 [Fig. 2]에서 볼 수 있듯이 2015년 글로벌 헬스케어 시장은 총 1조 5,759억달러 규모를 형성, 2016년은 전년대비 6.9% 성장하여 1조 6,844억 달러로 확대되었으며 헬스케어 IT분야는 2015년 시장 규모가 490억 달러 수준으로 아직 적은 비중을 차지하고 있으나 연 평균 8.1%의 성장률로 가장 빠른 확대가 전망되고 있다.



[Fig. 1] Global Healthcare Market Outlook

2016년 헬스케어 IT분야의 주요 이슈로는 병원과 연계한 가정용 어플개발, 가상/원격 모니터링, 밀레니얼²⁾ 유도전략 개발, 비용 조절을 위한 IT/BI 솔루션 사용, 헬스케어 데이터 관리 시스템 규제에 따른 수익성 보장요구 등이 시장형성에 주요한 영향을 미치고 있으며 고령화사회의 도래와 더불어 건강수명 및 삶의 질 향상을 위해 개개인이 적극적인 건강관리를 할 수 있도록 지원가능한 서비스의 개발에 관심이 집중되고 있다[6,7].

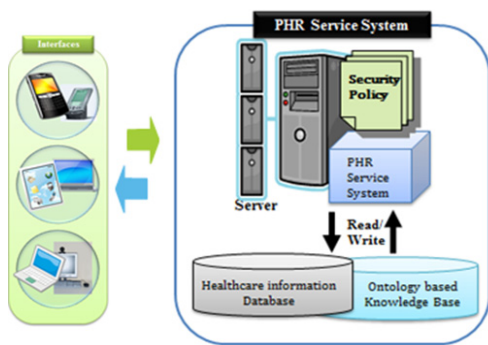


[Fig. 2] 2016 Market Share by Health Care

[Fig. 3]에서는 이러한 스마트 헬스케어 서비스 플랫폼인 지능형 PHR(Personal Health Record)서비스 시스템의 프레임워크를 보여주고 있다. 다수 사용자들의 다양한 개인건강기록(PHR: Personal Health Record)을 저장하고 PHR 프로파일을 분석하여 유사한 건강정보를 추출하고 이에 따라서 유용한 건강관리를 할 수 있도록 사용자별로 적절한 운동정보 및 식이정보 등을 추천해줄 수 있는 서비스를 제공해주는 프레임워크이다[8].

1) *4P : Personalized(개인화), Predictive(예측), Preventive(예방), Participatory(참여)

2) 밀레니얼: 디지털 언어와 장비를 태어나면서부터 사용함으로써 디지털적인 습성과 사고를 지닌 세대



[Fig. 3] Framework of Intelligent PHR Service System

본 논문에서는 PHR기반 맞춤형 건강정보 추천 서비스 플랫폼에서 지능형 건강정보 추천서비스를 제공할 수 있도록 온톨로지 기반 건강정보 모델링과 PHR 데이터 연관성 분석을 위한 통합알고리즘을 설계하고 시험데이터로 알고리즘의 유용성을 검증한다. 제안하는 알고리즘은 머신러닝 기법중 K근접이웃 알고리즘과 데이터마이닝 기법중 연관규칙 알고리즘을 확장한 알고리즘으로 사용자의 PHR 정보와 유사한 프로파일을 분석하고 최적의 건강정보를 추천하는 기능을 제공할 수 있다.

2. 관련연구

2.1 온톨로지 개념

온톨로지는 1967년 S. H. Mealy에 의해 전산학 분야에 도입되어 사용되었으며, 최근에는 인공지능(Artificial Intelligence) 분야에서 시멘틱 웹을 표현하는 중요한 언어로 각광받고 있다. 본래 온톨로지는 19세기 독일 철학자들에 의해 처음으로 사용된 단어로서 희랍어 ‘ontos(being)’와 ‘logos (word)’에서 유래하며, “세상의 어떤 관점을 세계에 존재(being)하는 것들의 종류, 그 본성과 관계 특성 등을 설명하는 분류체계(Taxonomy)를 제공하는 것”이라 정의된다[9,10].

온톨로지에 대하여 Sowa는 “온톨로지란 어떤 특정 영역(domain of interest)에서 존재하거나 존재할 수 있는 대상들의 범주(category)에 관한 연구”라고 정의하였으며 토마스 그루버는 온톨로지를 “공유하는(shared) 개념화(conceptualization) 형식적(formal)이고 명확한(explicit) 명세(specification)”라고 정의하였다[11]. 즉 합

의된 지식으로 몇몇의 개인이 임의로 정한 것이 아니라 해당 영역에 관련된 모든 구성원의 동의에 의해 수용되는 지식을 말한다.

온톨로지는 객관적으로 정해진 공통의 포맷을 이용하고 정보에 의미를 부여하는 것이 가능하므로 시스템이 상황 정보를 공유하고 이해 및 추론이 가능하다. 각각의 리소스들은 의미를 가지고 연결되며, 주어진 규칙에 따라서 규칙기반 추론이 가능하다. 이러한 점을 이용하여 사용자의 상황에 가장 알맞은 정보를 추천하여 제공할 수 있다. 이러한 의미를 가진 정보를 표현하는데 가장 대표적인 언어로 RDF, RDF Schema, OWL 등이 있으며, 이를 활용할 수 있도록 돕는 플랫폼은 Apache Jena, Pellet, FaCT 등이 있다[12]. 본 논문에서는 개인 맞춤형 건강정보 추천하고 사용자별로 추천을 위하여 먼저 건강정보의 명세 정의의 한 방법으로 온톨로지를 적용하고자 한다. 온톨로지를 적용하여 건강정보 명세를 정의하면 상황에 따른 건강정보 탐색을 위한 추론 규칙들을 유용하게 활용할 수 있을 것이다.

2.2 데이터마이닝

데이터마이닝은 존재하는 데이터로부터 과거에는 알지 못했던 새로운 데이터 모델이나 패턴을 발견하여 의사 결정에 사용하는 것이 지식탐사기술이다. 즉 데이터 속에 숨겨져 있던 관계들을 발견해 내는 것으로 데이터마이닝은 통계학에서 패턴인식에 이르는 다양한 기법을 사용하는 지식탐사기술이다[13,14].

대표적인 데이터마이닝 기법에는 분류, 군집화, 연관규칙 등이 있으며, 분류(Classification)는 특정 집단 내의 구성원들을 구분 및 분류한다. 군집화(Clustering)는 구성원들의 특성(속성)을 바탕으로 유사한 그룹을 생성한다. 연관성(Association)은 동시에 발생한 사건간의 관계를 정의한다. 연속성(Sequencing)은 특정 기간에 걸쳐 연속으로 발생한 사건간의 관계를 규명하며, 예측(Forecasting)은 현 데이터의 패턴을 바탕으로 미래의 요소를 예측하는 기술이다. 개인맞춤형 헬스케어서비스 플랫폼에서는 데이터마이닝 기법 중 분류에 속하는 K근접이웃 알고리즘을 적용하여 비슷한 유형의 PHR프로파일을 찾아 맞춤형 건강정보 추천에 적용할 수 있다[15, 16].

사용자 맞춤형 의료 데이터 관리 서비스 플랫폼은 병원에서 측정되는 기초 임상, 영상, 약물 처방 데이터와 스

마트 기기에서 측정 또는 사용자가 입력 가능한 음식물, 수면, 감정, 운동에 대한 데이터를 종합적으로 관리하며, 환자 동의하에 병원정보 시스템 및 특정 의료진에 자신의 건강 정보를 제공하거나 타인에게 제공하는 의료 서비스 플랫폼으로 개인 맞춤형 건강정보 제공 서비스를 위해서는 건강정보 온톨로지를 구축하고 PHR프로파일을 찾아 건강정보추천에 활용하기 위한 유사성 검사를 위해 데이터마이닝 알고리즘을 적용할 수 있다.

3. 건강정보 탐사 알고리즘 설계

3.1 온톨로지 기반 건강정보 모델링

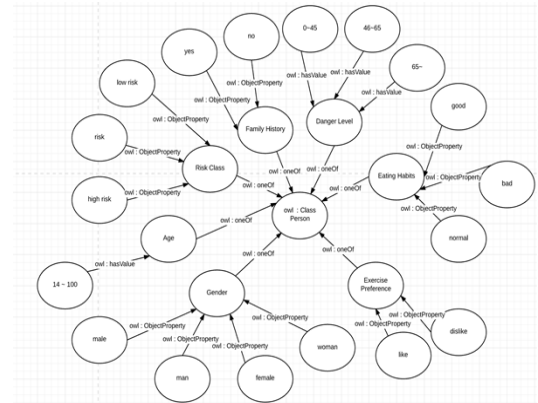
PHR기반 맞춤형 건강정보 추천 서비스 플랫폼인 바빌론에서 지능형 건강정보 추천서비스를 수행하기 위해서는 온톨로지 기반의 상황정보 모델을 구축하고 추론규칙을 정의하여 개인맞춤형 추천정보를 탐사할 수 있는 알고리즘을 적용하여 유사 profile을 분석하고 적절한 건강정보 제공 서비스를 제공하기 위해서 PHR정보를 온톨로지 기반의 상황정보 모델로 구성하였다. OWL(Web Ontology Language)을 이용하여 클래스와 속성 그리고 제약조건등을 정의하였다. 다음은 OWL의 각 구성요소의 개념들이며 그림 3은 OWL을 이용하여 PHR기반 건강정보를 온톨로지 기반의 상황정보로 모델링한 것을 보여준다.

【OWL 구성요소】

- owl:Class 어떤 개념을 추상화하여 표현한다. 모든 개체는 적어도 한 개의 owl:Class에 속하지 않으면 안된다.
- owl:oneOf OWL클래스의 인스턴스를 과부족 없이 열거한다. owl:one of속성의 정의역은 owl:Class이고, 치역은 rdf:List이다.
- owl:ObjectProperty 어떤 개체를 또다른 개체에 관련시키는 개체값형 속성이 속하는 클래스이다.
- owl:hasValue 주어가 되는 Restriction클래스의 모든 인스턴스들에 대해서, 해당 제약 대상인 속성 값들 중에서 적어도 1개가 목적으로 표현된 값과 같음을 나타낸다.
- owl:sameAs 2개의 개체가 (이름은 서로 다르더라도) 동일하다는 것을 나타낸다. 정의역과 치역은 모두 owl:Thing이다.

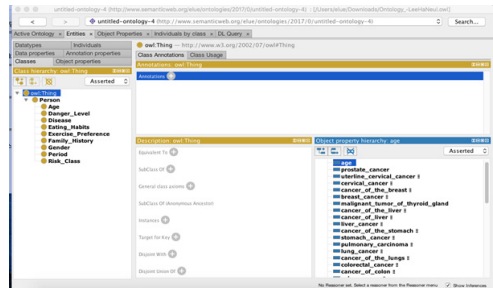
【OWL을 사용한 구조도】

처음 class를 Person이라는 개체로 정의했으며, 속성들은 Person클래스의 인스턴스로 열거하였다. 속성들은 성별, 나이, 운동 선호, 식습관, 가족력, 위험수준, 위험 등급 등이며 [Fig. 4]는 OWL로 표현된 구조도를 보여준다.

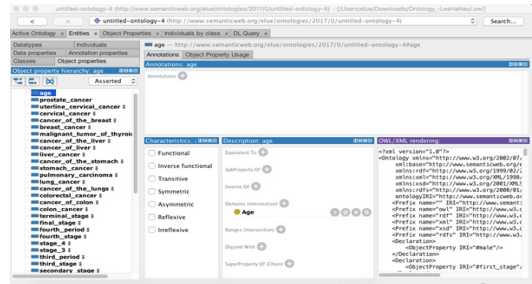


[Fig. 4] Ontology based Context Model using OWL

Protégé를 사용하면 온톨로지의 명세가 가능하며 [Fig. 5]-[Fig. 6]에서는 Protégé에서 OWL을 사용하여 ontology를 명세한 것을 보여주며 ontology.owl파일을 열어서 확인이 가능하다[17].



[Fig. 5] Ontology Class Hierarchy in Protégé

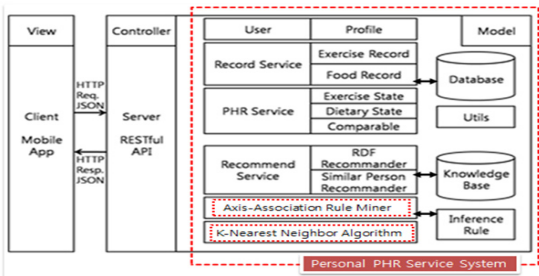


[Fig. 6] Object Property in Protégé

각각의 속성에는 이름은 다르지만 동일한 속성을 의미하는 Equivalent To를 정의하고 속한 범위를 뜻하는 Domains를 정해줄 수 있다.

3.2 알고리즘 설계

본 논문에서는 PHR기반 맞춤형 건강정보 추천 서비스 플랫폼에서 지능형 건강정보 추천서비스의 핵심 기술인 개인맞춤형 추천정보 추출 알고리즘의 설계하기 위해서 기존의 기계학습 알고리즘인 K근접이웃알고리즘과 데이터마이닝 알고리즘인 연관규칙 알고리즘을 확장하였다. 이는 사용자들의 PHR 프로파일을 분석하여 유사한 질환의 사용자들을 분류하여 유사 생체정보 및 건강정보를 비교하여 이를 기반으로 건강한 사용자의 건강행태에 따른 추천정보를 탐사하여 제공할 수 있도록 한다. [Fig. 7]은 PHR기반 맞춤형 건강정보 추천 서비스 플랫폼의 아키텍처를 보여준다.



[Fig. 7] PHR Service System Architecture

PHR 서비스 시스템의 구조에서 추천규칙을 탐사하기 위해서 K근접이웃 알고리즘으로 유사 프로파일 탐사를 수행한 후 속성기반 연관규칙 마이닝 알고리즘으로 연관 지식 탐사를 수행하여 추천규칙들을 생성하는 역할을 수행하게 된다. K근접이웃알고리즘은 유클리드의 거리를 계산하여 사용자의 프로파일의 유사도를 계산하게되며 유사도에 따라서 사용자의 그룹을 생성하게된다. 그룹의 개수는 조정이 가능하며 본 알고리즘의 실험을 위해서 6개의 그룹을 생성할 수 있도록 초기값을 설정하였다. [Fig. 8]은 제안하는 K근접이웃알고리즘의 수도코드이며 [Fig. 9]는 본 논문에서 제안하는 속성기반 연관규칙 알고리즘을 보여준다. 또한 속성 기반 알고리즘은 기존의 대표적인 연관규칙 알고리즘인 어프라이어리 알고

리즘을 확장하여 중요한 속성을 선택하고 선택된 속성을 기반으로 연관규칙을 수행하도록 설계하였다. 이는 불필요한 혹은 중요하지 않은 속성에 대한 후보 항목 집합의 생성을 차단하는 효과로 지식탐사의 효율성을 높일 수 있다.

```

1. classify
Classify(inX, dataset, labels, k){
    // inX:input vector, dataset: training matrix
    // labels: labeled vector, k : number of nearest kneighbor
    for i = 1 to m do
        compute_distance d(inX) //Euclidean Distance  $D = \sqrt{((xA_0 - xB_0)^2 - ((xA_1 - xB_1)^2)}$ 
    End for
        sort by ascending and store k shortest distances
    return most frequent item
}
2. ClassTest
ClassTest(){
    set hoRaio
    dataMat,dataLabels=file2matrix("filename")
    normMat, range, minVals = autoNorm(DataMat) //normalize
    compute number of vectors for test
    decide use type for test or for training
    m = normMat[0] // normMat[0] array size
    nv = number of vector for test
    errorCount = 0.0
    for i=0 to nv
        result = classify(normMat[i to end],normMat[nv to m],DataLabels[nv to m],3)
        if(result != dataLabels[i]):
            errorCount += 1.0
    print errorCount
}
    
```

[Fig. 8] K-Nearest Neighbor Algorithm

```

Axis-based Apriori()
{
    // T Set of transactions t
    // Ck Candidate set size k
    // Lk Frequent itemset size k

    Select_attribute()
    L1 ← {frequent items}
    k ← 2
    while Lk-1 ≠ ∅
        Ck ← {a ∪ {b} | a ∈ Lk-1 ∧ b ∉ a} - {c | {s | s ⊆ C ∧ |s| = k - 1 ∉ Lk-1}
        for transactions t ∈ T
            Ct ← {c | c ∈ Ck ∧ c ⊆ t}
            for candidates c ∈ Ct
                count[c] ← count[c] + 1
            Lk ← {c | c ∈ Ck ∧ count[c] ≥ ε}
        k ← k + 1
    return ∪k Lk
}
    
```

[Fig. 9] Axis-based Apriori Algorithm

4. 실험 및 평가

4.1 테스트 데이터 수집 및 전처리

먼저 건강정보 모델링과 기계학습, 데이터마이닝 알고리즘 등의 테스트 수행을 위해서 바빌론 광장에서 요구되는 데이터요구사항 명세를 정의하였다. 테스트 데이터 구축은 랜덤방식으로 일반인들의 데이터와 건강질환자들의 데이터들로 구성하였으며 알고리즘의 성능 측정을 위하여 테스트 데이터를 임의 생성하여 실험을 진행하였다. 테스트데이터들은 일반적인 수치데이터와 범위값으로 이루어져 있어 전처리가 필요하며 식습관, 운동선호도, 가족력, 질병단계 등의 데이터 값을 이산값으로 변환하여 실험 및 평가를 진행하였다.

테스트 데이터는 일반인과 암환자 PHR정보 1000개의 데이터를 수집한 데이터와 많은 양의 데이터 비교를 위한 10000개의 튜플을 가지는 테스트데이터셋을 전처리하여 실험을 진행하였다. 임의의 규칙없는 데이터셋은 dataset A로 구성되며 1000개의 튜플을 가진다. dataset B는 연관규칙이 존재하는 테스트데이터로 구성하여 전처리를 수행하였으며 알고리즘 테스트를 위해서 유사한 랜덤데이터 9000여개를 추가하여 dataset C를 구성하여 1000개 그룹과 10000개의 튜플을 가지는 테스트 데이터셋으로 나누어 실험을 하였다. [Fig. 10]은 테스트데이터의 일부 보여준다.

| 이름 | 성별 | 나이 | 운동선호 | 식습관 | 질병 | 질병 단계 | 가족력 | 대상암 | 1 대상 : 암환자 |
|------------|--------|----|---------|--------|----|-------|-----|-------|------------|
| 1031474744 | male | 32 | like | normal | 8 | 1 | 0 | 폐암 | 2 |
| 1023547785 | female | 27 | dislike | bad | 7 | 2 | 1 | 위암 | 3 |
| 1032654578 | male | 47 | like | good | 8 | 3 | 1 | 간암 | 4 |
| 1023654731 | male | 28 | dislike | normal | 8 | 1 | 0 | 갑상선암 | 5 |
| 1064758137 | female | 35 | like | normal | 7 | 2 | 1 | 유방암 | 6 |
| 1031757842 | female | 48 | dislike | good | 7 | 4 | 1 | 자궁경부암 | 7 |
| 1047571237 | male | 53 | like | bad | 8 | 4 | 0 | 전립선암 | 8 |
| 1031758476 | female | 55 | like | good | 7 | 3 | 0 | | |
| 1067425871 | female | 35 | dislike | normal | 7 | 3 | 1 | | |
| 1032787456 | male | 40 | like | bad | 8 | 2 | 1 | | |
| 1085674215 | female | 57 | dislike | normal | 8 | 4 | | | |
| 1034752186 | female | 50 | like | good | 7 | | | | |
| **23654731 | male | 35 | like | normal | 8 | | | | |
| - | - | 23 | dislike | bad | - | | | | |

[Fig. 10] Test Dataset

4.2 실험

다양한 테스트 데이터 셋에 대한 실험을 수행하기 위해 테스트 데이터 개수에 변화를 주어 실험시켰으며 이를 통해 질병 단계에 따른 환자의 운동선호와 식습관을 유추 할 수 있어 콘텐츠를 유용한 정보를 제공하는 추천 서비스에 적용될 수 있다는 것을 확인할 수 있었다.

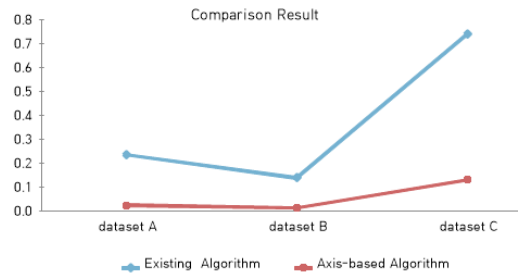
본 논문에서 제안한 속성기반 연관규칙 알고리즘의 경우 추가적으로 속성에 질병을 추가하면 질병과 질병 단계에 따른 환자의 상태를 유추해 콘텐츠를 추천할 수 있으며 추천 가능한 정보탐사에 효율적인 탐색시간을 제공해주는 것을 실험을 통해 확인하였다.

<Table 3> Experimental Result (unit:second)

| TestDataset category | dataset A | dataset B | dataset C |
|-------------------------|--------------|--------------|--------------|
| Existing Algorithm | 0.235 | 0.14 | 0.74 |
| Axis-based Algorithm | 0.023 | 0.013 | 0.13 |

<Table 3>은 기존의 연관규칙 알고리즘과 속성기반의 확장된 연관규칙 알고리즘을 각각의 테스트데이터셋에 대하여 실행시킨 결과를 보여주는 표이다.

특히 속성기반 연관규칙이 데이터셋 B와 C에 대해서 튜플수에 관계 없이 탁월한 성능을 보여주고 있음을 알 수 있다.



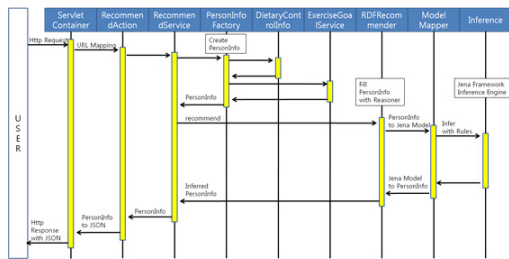
[Fig. 11] Comparison Result

[Fig. 11]은 실험결과를 그래프로 나타낸 것으로 표와 그래프에서 확인할 수 있듯이 제안하는 속성기반 연관규칙 알고리즘의 성능이 훨씬 뛰어나다는 것을 알 수 있으며 이는 특정 중요속성을 대상으로 연관규칙을 수행하는 것은 후보항목집합의 생성을 최소화할 수 있어서 가능한 것이다.

4.3 추천 알고리즘의 적용

PHR 기반 헬스케어 통합 서비스 플랫폼의 지능화 핵심 기능인 바빌론 광장의 추천시스템은 PHR 관리 저장소, 공공데이터 등의 빅데이터 분석, 사용자 프로파일 분

석에 따른 상황 정보 모델링, 집단 지성, 사용자 맞춤형 건강정보 및 콘텐츠 추천 등의 기능을 수행할 수 있도록 구축된다. 본 논문에서 제안하는 알고리즘은 사용자 상황인지 및 프로파일 분석에 따른 정보 및 콘텐츠 추천 및 추론시스템의 [Fig. 12]와 같은 기본적인 워크플로우에 적용되어 개인 맞춤형 건강정보를 추론하는 기능을 수행하며 추론규칙들은 지식베이스에 점진적으로 저장된다. [Fig. 12]는 PHR 서비스 시스템에서의 추론프로세스의 워크플로우를 나타낸 것이다.



[Fig. 12] Inference Workflow

5. 결론

본 논문에서는 PHR 기반 헬스케어 서비스 플랫폼 지능화를 위한 온톨로지 기반 건강 정보 모델 구축과 기계학습과 데이터마이닝 기법을 적용한 건강정보 추천 알고리즘을 설계하였다. 기존의 데이터마이닝 기법중 연관규칙 알고리즘을 확장하여 속성을 기반으로 연관규칙 탐색을 수행하여 지식탐사의 연관성을 높이고 효율적인 탐색 시간을 제공할 수 있도록 하였다.

사용자 프로파일별 사용자 그룹화를 하여 유사패턴의 사용자 프로파일을 검색하는 기능과 일상생활 전반에서 발생하는 행위 사건들의 발생빈도, 연관관계, 순차 및 반복 관계 등에 대한 데이터 분석기술을 적용하여 사용자의 질환과 건강상태에 따른 맞춤정보 생성이 가능하도록 한다. 특정 속성의 가중치를 주어 속성기반 연관규칙 알고리즘을 설계하여 기존의 어프라이어리 알고리즘보다 성능이 우수하다는 것을 실험을 통하여 확인할 수 있었다. 제안하는 알고리즘은 맞춤형 건강정보 추천서비스를 위한 개인맞춤형 헬스케어플랫폼에서 맞춤 정보 추론프로세스에 적용되어 개인화 추천서비스 제공에 활용되어

질 수 있다. 향후 제안 알고리즘의 신뢰도와 지지도에 따른 연관규칙 생성결과를 분석하여 개인맞춤형헬스케어 서비스 플랫폼에서 활용할 수 있는 최적의 신뢰도와 지지도 값을 찾는 연구가 필요하다.

ACKNOWLEDGMENTS

This paper was supported by Konkuk University.

REFERENCES

- [1] J.Y.Choi, "Healthcare 3.0 : A world of healthy life", CEO information Vol.831, Samsung Economic Research Institute, 2011,
- [2] K.T.Oh, J.E.Lee, "Smart Life revolutionary real and smartphone addiction", Internet and Information Security, Vol.3 No.4, 2012.
- [3] J.H.Park, T.K.Hwangbo, "Healthcare IT Convergence Technology, Journal of Korea Information and Communications, Vol.28, No 5, 2011.
- [4] Donghyun Kim, Seoksoo Kim, "Design of Key Tree-based Management Scheme for Healthcare Information Exchange in Convergent u-Healthcare Service ", Journal of the Korea Convergence Society, Vol. 6, No. 6, pp. 81-86, 2015.
- [5] E. Choi, S. Lee, "Access Control Mechanism based on MAC for Cloud Convergence", Journal of the Korea Convergence Society, Vol. 7 No. 1, pp. 1-8, 2016.
- [6] J.W.Song, S.H.Kim, M.E.Jung, "Protection of health information for u-healthcare service", Journal of Information Security, Vol.17, No 1, 2007.
- [7] S.H.Kim, "Health-IT convergence technology trends and forecasts (mainly U-health)", Journal of Electronics, Vol.37, No6, 2010.
- [8] M. Shin, H.Jeon and B.Lee, "Constructing RBAC based Security Model in u-Health Care Service Platform", SWJ(937914), 2014.
- [9] L.Kim, "Convergence of Information Technology

- and Corporate Strategy”, Journal of the Korea Convergence Society, Vol. 6, No. 6, pp. 17-26, 2015.
- [10] H.Chung, J. Kim, “Design of Semantic Models for Teaching and Learning based on Convergence of Ontology Technology”, Journal of the Korea Convergence Society, Vol. 6, No. 3, pp. 127-134, 2015.
- [11] Young Sung Cho and Song Chul Moon, “Weighted Mining Frequent Pattern based Customer’s RFM Score for Personalized u-Commerce Recommendation System”, Journal of Convergence, Vol.4. No.4. Dec. 30, 2013.
- [12] You-Dong Yun, “Development of Smart Senior Classification Model based on Activity Profile Using Machine Learning Method”, Journal of the Korea Convergence Society, Vol. 8. No. 1, pp. 25-34, 2017.
- [13] Jang, Kyoung Hee, Jung, In Sook, “Converged Study on the Nurses’ Knowledge and Performance of Cancer Pain Management in one city”, Journal of the Korea Convergence Society, Vol. 7. No. 6, pp. 115-124, 2016.
- [14] Menaouer Brahami, Baghdad Atmani and Nada Matta, “Dynamic knowledge mapping guided by data mining: Application on Healthcare”, Journal of Information Processing Systems, Vol.9, No.1, 2013.
- [15] Reid, I. Cheong, M. Henrickson, and J. Smith, “A Novel Use of RBAC to Protect Privacy in Distributed Health Care Information Systems”, ACM Transactions on Information and System Security, 2014.
- [16] Byung-Jin Jeon, Deok-Byeong Yoon, Seung-Soo Shin, “Integrated Monitoring System using Log Data”, Journal of IT Convergence Society for SMB, 2017.
- [17] Jin-Hee Ku, “A Study on the Machine Learning Model for Product Faulty Prediction in Internet of Things Environment”, Journal of IT Convergence Society for SMB, 2016.

신 문 선(Shin, Moon Sun)



- 2004년 8월 : 충북대학교 전자계산학과(이학박사)
- 2005년 8월 ~ 현재 : 건국대학교 컴퓨터공학과 교수
- 관심분야 : ICT 융복합기술, 상황인지 및 추론, 빅데이터 분석, 헬스케어
- E-Mail : msshin@kku.ac.kr