

연관분석과 순차분석을 통한 스마트홈 패턴 구성방안

정승민¹, 최한얼¹, 광경호¹, 김민재¹, 정해린¹
¹울산대학교 IT융합전공 학부생

wjddudflal@naver.com, ulywooly@gmail.com, gyeonghogwaq@naver.com,
us2922142@gmail.com, jhr2996@naver.com

A Study on the Composition of Smart Home Patterns through Association Analysis and Sequential Analysis

Seung-Min Jeong¹, Han-Eol Choi¹, Gyeong-Ho Gwag¹,
 Min-Jae Kim¹, Hae-Rin Jeong¹
¹Dept. of IT Convergence, Ulsan University

요 약

스마트홈은 기술 시스템, 자동화 프로세스, 원격 제어 기기 등을 아파트나 주택에서 사용하는 것을 말한다. 주요 목적은 가정에서 삶의 질과 편의성을 높이는 것이다. 현재의 스마트홈은 사용자의 원격 제어 방식을 사용하고 있다. 이러한 방식은 고정된 시간에만 스마트홈이 작동하도록 한다는 문제가 있었다. 연관분석과 순차분석을 통해 AI가 상황과 사용자의 취향을 학습한다면, 스스로 최적화된 패턴을 제공할 수 있을 것이다.

1. 서론

스마트홈은 가전제품(TV, 에어컨, 냉장고 등)을 비롯해 에너지 소비장치(수도, 전기, 냉난방 등), 보안기기(도어록, 감시카메라 등) 등 다양한 분야에서 모든 것을 통신망으로 연결해 모니터링, 제어할 수 있는 기술을 말한다. 스마트폰이나 인공지능(AI) 스피커가 사용자의 음성을 인식해 집 안의 모든 사물 인터넷(IoT) 기기를 연결하고 사용자의 특성에 따라 자동으로 작동하거나 원격으로 조종할 수 있다 [1][2]. 현재의 스마트홈은 원격제어의 방식을 주로 사용하고 있다. 우리는 더 나아가 AI가 상황과 사용자의 취향을 학습하여 스스로 패턴을 제공하는 방법을 연구하였다.

2. 아이디어

2.1. 연관분석

연관분석은 항목 간의 일정한 연관 규칙을 발견하는 분석이다. 연관분석의 측정지표로는 지지도(support), 신뢰도(confidence), 향상도(lift)가 있다. 지지도는 조건절이 일어날 확률로, A와 B가 동시에 발생할 확률이다. 신뢰도는 조건절이 주어졌을 때,

결과절이 일어날 조건부 확률을 의미하고, 향상도는 조건절과 결과절이 서로 독립일 때와 비교해 두 사건이 동시 발생하는 확률이다.

지지도	$P(A \cap B) = \frac{A와 B가 동시에 발생한 수}{전체 발생 수}$
신뢰도	$\frac{P(A \cap B)}{P(A)} = \frac{A와 B가 동시에 발생한 수}{A가 발생한 수}$
향상도	$\frac{P(B A)}{P(B)} = \frac{P(A \cap B)}{P(A)P(B)} = \frac{A와 B가 동시에 발생한 수}{A가 발생한 수 * B가 발생한 수}$

(그림 1) 지지도, 신뢰도, 향상도

스마트홈에 사용되는 디바이스들은 통신망을 통해 플랫폼과 연결되어 있다. 이때, 디바이스들을 사용하면서 플랫폼에는 해당 디바이스의 작동(전원, 작동시간, 상태 등)에 대한 데이터가 전달된다. 여기서 전원의 상태변화에 대한 데이터를 사용하여 트랜잭션을 생성한다. 플랫폼을 통해 제어되는 디바이스들의 전원이 모두 off로 설정된 1번 리스트를 만들고, 이후 디바이스의 전원에 변화가 생길 때마다 해당 변화를 반영한 리스트를 추가한다. 이 과정을 거쳐 만들어진 트랜잭션들을 통해 연관분석을 실행하면 어떤 디바이스들이 동시에 동작하는지에 대해 알 수 있다.

```

Step 1 : [1_D_off, 2_D_off, 3_D_off, 4_D_off, 5_D_off] # 기본 트랜잭션
Step 2 : [1_D_on, 2_D_off, 3_D_off, 4_D_off, 5_D_off] # 1번 디바이스 on
Step 3 : [1_D_on, 2_D_off, 3_D_off, 4_D_on, 5_D_off] # 4번 디바이스 on
Step 4 : [1_D_on, 2_D_off, 3_D_off, 4_D_on, 5_D_on] # 5번 디바이스 on
Step 5 : [1_D_on, 2_D_off, 3_D_off, 4_D_off, 5_D_on] # 4번 디바이스 off
    
```

(그림 2) 트랜잭션 생성 예시

분석 과정에서 최소 지지도, 최소 신뢰도, 최소 향상도를 조절하여 자주 나타나는 패턴 항목을 찾을 수 있다.

2.2. 순차분석

순차분석은 시간과 순서를 고려한 분석이다. 앞의 연관분석에서는 자주 등장하는 패턴을 찾는 것은 가능했지만, 해당 패턴이 어떤 시간에, 어떤 순서로 발생하는지는 알 수 없었다. 그렇기에 시계열 데이터를 반영하는 순차분석을 추가적으로 실시하였다. 플랫폼에서 전원의 상태변화가 일어나는 경우의 데이터만으로 리스트를 생성하고, 해당 상태변화가 언제 발생하는지에 대응하는 리스트들을 생성한다.

```

# 상태 변화 항목
['9_D_on', '15_D_on', '4_D_on', '4_D_off', '18_D_on', '18_D_off', '15_D_off', '16_D_off',
'10_D_on', '10_D_off', '10_D_on', '10_D_off', '9_D_off', '1_D_on', '22_D_on']

# 상태변화가 일어난 시간
['00:00:05', '00:01:28', '00:25:28', '00:26:35', '00:36:49', '00:38:05', '00:39:35',
'00:39:35', '00:57:10', '00:59:05', '01:01:47', '01:03:35', '01:04:00', '07:00:15', '07:00:44']
    
```

(그림 3) 상태변화 항목과 대응 시간 리스트 예시

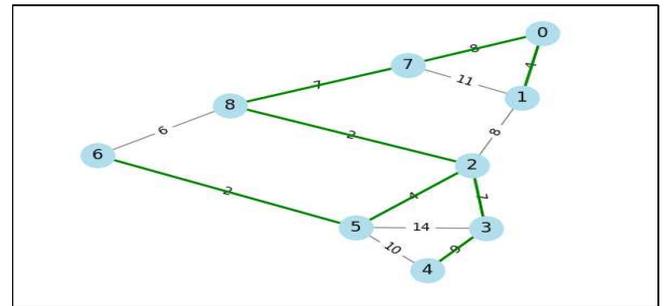
리스트 하나의 단위는 24시간으로 하여, 각 요일과 오전*오후에 해당하는 리스트들이 중첩리스트를 이루게 하였다.

중첩리스트에 대해 패턴이 발생하는 최소 반복 횟수와 각 상태변화가 일어나는 사이의 최대 간격인 gap[3]과, 패턴의 시작과 끝으로 발생할 수 있는 최대 시간인 span[3]을 지정하여 패턴을 탐색한다. 이를 통해 특정 순서로, 특정 시간에 반복되는 패턴의 리스트와 해당 패턴이 나타나는 횟수를 발견할 수 있다.

스마트홈을 제어하는 플랫폼의 AI는 사용기간이 증가할수록 사용자의 데이터를 학습해 나갈 것이다. 연관분석과 순차분석에서 공통적으로 나타나는 패턴 항목들은 규칙성을 가지게 될 것이고, AI는 해당 지표들을 사용하여 스스로 패턴을 제공하는 것이 가능할 것으로 기대된다.

3. 스마트홈 패턴 구성방안의 향후 과제

스마트홈 패턴을 구성하는 과정에서 시간적 요소만을 고려해야 하는 것이 아니다. 예를 들어, 거실과 테라스나 침실과 드레스룸처럼 연관된 장소에 있거나 디바이스끼리의 거리가 가까울 경우 화장실과 발코니 같은 연관성 없는 장소의 패턴보다 더 중요성을 가질 것이다. 더 정확한 패턴을 출력하기 위해서는 디바이스의 특성과 디바이스 사이의 거리에 대한 가중치를 부여할 필요가 있다. 이런 패턴에 가중치, 혹은 제약조건을 추가해 나간다면 사용자에게 최적화된 스마트홈 패턴을 제공할 수 있을 것이다.



(그림 4) 엣지를 통해 가중치를 표현한 예시[4]

참고문헌

[1] 네이버 시사상식사전
 [2] 홍석일, 스마트 홈 기술 동향, 한국통신학회지, 제37권, 제11호, 28-35p, 2020
 [3] Xin Wang, Seq2Pat: Sequence-to-Pattern Generation for Constraint-Based Sequential Pattern Mining, aai-22 vol.36 No.11, 2022
 [4] NetworkX documentation
https://networkx.org/documentation/stable/auto_examples/graph/plot_mst.html