유비쿼터스 홈 네트워크 시스템에서 은닉 마르코프 모델을 이용한 사용자 행동 상태 분석 및 예측 알고리즘☆

Analysis and Prediction Algorithms on the State of User's Action Using the Hidden Markov Model in a Ubiquitous Home Network System

신 동 규* 신 동 일** 황 구 연*** 최 진 욱****
Dong-kyoo Shin Dong-il Shin Gu-youn Hwang Jin-wook Choi

요 약

본 논문은 유비쿼터스 홈 네트워크 시스템에서 저장된 사용자 행동 프로파일 데이터에 은닉 마르코프 모델에 적용하여 사용자의 행동 상태를 예측하는 알고리즘을 제안한다. 은닉 마르코프 모델은, 순차 데이터를 갖는 패턴을 인식하기 위해서 데이터에 내포되어 있는 시간성을 적절히 표현하고, 그것으로부터 원하는 정보를 추론할 수 있는 대표적인 모델이다. 제안 알고리즘에서는 "행동 인지 시스템(Activity Recognition System)"에 의하여 저장된 행동 발생 횟수, 행동 지속시간, 행동이 발생된 위치 데이터를 학습 데이터로 이용하였다. 사용자의 행동에 가중치를 부여하여 사용자의 행동에 대한 흥미를 객관적으로 수식화 하는 방법을 제안하였으며 은닉 마르코프 모델을 이용하여 시간에 따른 가중치 변화를 구하여 사용자의 행동 상태 변화를 예측하였다. 제안 알고리즘은 현실적인 유비쿼터스 홈 네트워크 구축에 도움을 준다.

ABSTRACT

This paper proposes an algorithm that predicts the state of user's next actions, exploiting the HMM (Hidden Markov Model) on user profile data stored in the ubiquitous home network. The HMM, recognizes patterns of sequential data, adequately represents the temporal property implicated in the data, and is a typical model that can infer information from the sequential data. The proposed algorithm uses the number of the user's action performed, the location and duration of the actions saved by "Activity Recognition System" as training data. An objective formulation for the user's interest in his action is proposed by giving weight on his action, and change on the state of his next action is predicted by obtaining the change on the weight according to the flow of time using the HMM. The proposed algorithm, helps constructing realistic ubiquitous home networks.

☞ keyword : 유비쿼터스 홈 네트워크 (Ubiquitous Home Network), 은닉 마르코프 모델 (Hidden Markov Model), 사용자 행동 예측 (Prediction of User's Activity), 데이터 마이닝 (Data Mining)

1. 서 론

유비쿼터스 기술은 인간의 생활에 급속하게 스며들고 있고 주택에도 유비쿼터스 홈 기술이 적용되고 있다. 유비쿼터스 홈은 유무선 통신 네트워크를 기반으로 가정 내의 다양한 가전 기기 및 센서들이 네트워크로 상호 연결되어 다양한 서비스 제공이 가능하며, 또한 능동적이고 상황적응적인 상호 연동을 통한 사용자의 편의를 극대화시킨다. 즉, 삶의 질 향상을 추구하는 지능화 된 가정 내의 생활환경이라 정의할 수 있다[1].

거주 공간을 제어하고 거주자의 의도를 파악하

* 정 회 원 : 세종대학교 컴퓨터공학과 교수
shindk@sejong.ac.kr
** 정 회 원 : 세종대학교 컴퓨터공학과 교수
dshin@sejong.ac.kr
*** 준 회 원 : 세종대학교 대학원 석사과정
hgy1999@gce.sejong.ac.kr
**** 준 회 원 : 세종대학교 대학원 석사과정
seekpeace@gce.sejong.ac.kr
[2010/11/09 투고 - 2010/11/18 심사 - 2011/02/11 심사완료]
☆ 본 논문은 지식경제부의 산업원천기술개발사업(과제번호: 10011346)으로 지원된 연구임.
☆ 본 논문은 창립10주년 2010년도 한국인터넷정보학회 학술발표대회 우수논문의 확장버전임.

여 상황에 적합한 서비스를 제공하는 연구는 상당부분 진행되어 있다. AwareHome, EasyLiving, Intelligent Space 그리고 SensingRoom은 거주하는 공간에서 센서들을 사용하여 편의성을 높여주는 시스템의 대표적인 예이다[2-6].

유비쿼터스 기술이 발전함에 따라 유비쿼터스 환경에서 사용자의 편의를 도모하기 위해 생활습관과 움직임에 맞추어 행동을 예측 하는 것이 중요한 요구 사항으로 대두되고 있다.

본 논문의 목적은 사용자 행동 프로파일 데이터에 은닉 마르코프 모델(Hidden Markov Model - HMM)을 적용하여 사용자의 행동 상태를 예측하고 이를 통하여 최적의 서비스를 제공하는 것이다. 본 연구의 실험 방법으로는, “MIT Placelab”에서 “행동 인지 시스템(Activity Recognition System)[7]” 구현 결과로 측정된 행동 프로파일 데이터를 HMM 도구에 적용을 하여 행동 상태 예측 정확도를 구한다. 행동상태를 정의하기 위해 사용자의 행동에 가중치를 부여 하였으며, 행동에 대한 흥미를 객관적으로 측정 할 수 있는 가중치를 수식화 하는 계산식과 행동 프로파일 데이터에 HMM을 적용하는 방법을 알고리즘으로서 제안한다.

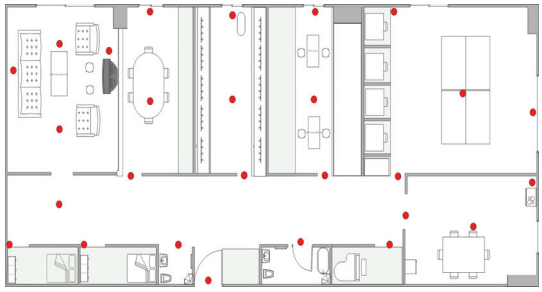
본 논문의 구성은 다음과 같다. 2장에서는 행동예측 시스템에 관한 관련연구를 살펴보고, 3장에서는 HMM을 기술하며, 4장에서는 HMM 기반 행동 상태 예측 방법을 살펴본다. 5장에서는 실험 및 결과를 설명하며, 6장에서는 결론 및 향후 연구방향에 대해서 논한다.

2. 관련연구

유비쿼터스 환경에서 사용자의 상황 공간을 설명하며 데이터 마이닝 기술을 이용한 행동 순차 패턴 발견에 대해서 기술한다.

2.1 유비쿼터스 환경에서 마이닝을 이용한 행동순차 패턴

유비쿼터스화는 사물에 소형 컴퓨터와 같은 지



(그림 1) 유비쿼터스 홈에서 상황 공간

능을 부여하고 언제 어디서나 접속이 가능한 네트워크를 통해 지능화된 사물의 정보를 전자공간의 정보와 실시간 공유함으로써 자체적으로 진단, 분석, 판단, 조치를 가능하도록 한다는 개념이다. 유비쿼터스 환경에서는 사용자의 현재 위치, 행동, 감정 등을 객체로 나타낼 수 있으며, 사용자의 객체에 대한 정보 값과 그 정보들의 변화를 상황(Context)이라고 정의한다. 이러한 상황 정보를 유비쿼터스 홈에서 얻어내거나, 상황에서 객체의 활동에 따라 발생되는 정보를 얻기 위해 상황 공간을 구성해야 한다[8]. 유비쿼터스 홈에서는 적당한 태그의 수로 위치 측정이 이루어져야 하며, 태그를 이용한 행동 패턴발견을 위한 상황 공간을 (그림 1)에 나타내었다. 빨간 점은 서비스를 제공하는 위치를 의미한다. 사용자, 즉 이동객체가 현관문을 통해 유비쿼터스 홈 공간에 진입하여 태그를 통하여 초기 위치를 설정한 후 이동하면서 데이터 및 태그의 ID를 수집한다. 수집한 순차 데이터는 위치 서버로 전송되게 된다[9].

순차 데이터로부터 의미 있는 지식을 탐사하기 위한 순차 패턴 데이터마이닝 기법에 대한 여러 연구가 진행되었다[10]. 순차 패턴 데이터마이닝은 방대한 양의 누적된 서로 상이한 시간 데이터로부터 이전에 잘 알려지지는 않았지만, 묵시적이고 잠재적으로 유용한 시간 지식을 추출하는 기술로 정의된다. 현재까지 시간 데이터로부터 패턴 지식을 탐사하기 위한 순차 패턴 데이터마이닝은 크게 시간규칙 탐사, 시퀀스마이닝, 경향분석 등으로 분류된다. 시간규칙탐사 기법은 기존 데이터

(표 1) HMM의 5가지 요소

1. 은닉 상태 집합(hidden state set) : 마르코프 프로세스에 의해서 설명되는 상태들의 집합
2. 관찰가능 상태 집합(observable state set) : 외형적으로 눈에 보이는 전이 상태들의 집합
3. π 벡터 : 특정 은닉 상태가 시간 t=1일 때 모델의 확률
4. 상태전이 행렬 : 이전의 은닉 상태에서 현재의 은닉 상태로의 전이 확률을 나타내는 것으로 모델 내부의 은닉 상태들 간의 전이 확률을 나타내는 확률
5. 관찰확률 행렬 : 특정 은닉 상태에서의 관찰 가능 한 한 각각의 상태들에 대한 확률을 나타내는 행렬

마이닝 기법에서 정의한 규칙을 확장하여 시간관계 및 상호연관 관계를 포함한 시간규칙을 찾는 기법으로, 여기에 속한 기법으로는 주기적으로 반복되는 연관규칙을 탐사하는 주기적 연관 규칙 탐사, 달력으로 표현된 시간 패턴을 가지는 연관규칙을 탐사하는 캘린더 연관규칙 탐사 등이 있다.

3. Hidden Markov Model

HMM은 시간적 변화를 가지는 데이터의 인식, 음성 인식, 분류(Classification) 분야에 주로 사용된다. HMM은 관찰이 불가능한 은닉(hidden)의 확률적 과정(stochastic process)을 관찰이 가능한 기호(symbol)를 발생시키는 다른 확률적 과정을 통하여 모형화(modeling)하는 이중의 확률론적 과정이다[11]. 행동 상태 예측 시스템 자체가 사용자의 의도가 은닉된 상태에서 사용자의 행동에 관한 흥미의 변화도를 감지하는 문제로 파악 될 수 있기 때문에 댁 내 의 거주자의 행동 상태 예측 분야에도 효과적으로 적용될 수 있다.

HMM은 (표 1)과 같이 2개의 상태 집합과 3개의 확률 집합으로 구성되는 5개의 요소를 갖는다. 이와 같이, HMM은 관찰 가능한 상태들과 은닉 상태들 간의 확률적 관계를 이용하여 계산되는 알고리즘이다.

HMM은 기호로 $\lambda=(\pi, A, B)$로 표현되며 은닉 상태수를 N, 관측 심벌의 개수를 M, 관측 순서의 길이를 T라고 할 때 다음과 같은 구성요소들을 가진다.

- $Q=\{q_1, q_2, q_3..., q_n\}$: 은닉상태 집합
- $V=\{v_1, v_2, v_3,..., v_n\}$: 각 상태마다 관측 가능한 심벌의 집합
- $O=\{o_1, o_2, o_3,..., o_n\}$: 관찰 심벌로 이루어진 관측열
- $\pi=\{\pi_i\}$, $\pi_i = P(q_i \; at \; t=1)$: 모든 상태의 초기 확률 분포
- $A=a_{ij}$, $a_{ij} = P(a_j \; at \; t | q_i \; at \; t-1), 1 \le i,j \le N$: A는 상태 전이 확률 분포이며, a_{ij}는 상태 i에서 상태 j로 전이 확률을 나타낸다.
- $B=b_j(k), b_j(k) = P(v_k \; at \; t | q_j \; at \; t), 1 \le k \le M$: B는 관측 심벌 확률 분포이며, b_{jk}는 상태 j에서 v_k를 생성하는 확률을 나타낸다.

이렇게 정의된 HMM은 실제로 모델을 사용하기 위해서는 다음과 같이 세 가지 중요한 문제를 해결해야한다.

- 확률 추정(Probability Estimation) : 관측열 $O=\{o_1, o_2, o_3,..., o_{t-1}, o_t\}$, 모델 $\lambda=(\pi, A, B)$이 주어지면, 모델에 의해 생성되는 관측열에 대한 확률 $P(O|\lambda)$을 계산하는 과정
- 최적 상태 순서(Optimal Sequence) 결정 : 관측열 $O=\{o_1, o_2, o_3,..., o_{t-1}, o_t\}$가 주어지면, 최적화 기준에 따라 가장 가까운 상태열 $O=\{q_1, q_2, q_3..., q_n\}$을 구하는 과정
- 매개변수 추정(Parameter Estimation) : 관측열 $O=\{o_1, o_2, o_3,..., o_{t-1}, o_t\}$에서 $P(O|\lambda)$을 최대화하는 모델 $\lambda=(\pi, A, B)$을 조정하는 과정(관측들이 어떻게 출현되는 가를 가장 잘 표현하도록 모델 변수들을 최적화 시키는 문제)

본 논문에서는 거주자의 행동상태를 예측하는 문제에 두 번째 문제 해결 방법을 적용한다. 이 방법에서는 동적 프로그래밍 기법중의 하나인 Viterbi 알고리즘을 사용한다. Viterbi 알고리즘은 행동 인지 감지 센서에서 저장된 행동 특징 벡터

(표 2) State-change sensors와 ESM(Experience Sampling Method)을 이용하여 얻은 Data type

Activity[a]	Sensor ID[b]	Day[c]	Activation Time[d]	Deactivation Time[e]	Duration(s)[f]	Room(opt)[g]	Object Type(opt)[h]
Preparing Breakfast	PDA	12/01/02	08:23:01		10min		
	23	12/01/02	08:23:01	08:23:07	6	kitchen	drawer
	18	12/01/02	08:23:09	08:23:17	8	kitchen	cabinet
	89	12/01/02	08:24:49	08:24:59	10	kitchen	fridge door
			⋮				
			(many read-ings)				

a Attribute of the user's activity saved in the PDA
* b ID number of the state-change sensor
c Day when user's activity was sensed
d Sensor activation time in seconds starting from 12:00am
e Deactivation time in seconds starting from 12:00am
f Time that the sensor was activated
g Location of the sensor (room) in the house
h Household object in which the sensor was installed

와 훈련된 HMM을 기반으로 가장 확률이 높은 행동 상태열을 제공한다.

Viterbi알고리즘은 다음과 같다.

Ⅰ. 초기화

$$\delta_1(i) = \pi_i b_j(o_1) \ 1 \le i \le N \ .....(1.1)$$
$$\Psi_1(i) = 0$$

Ⅱ. 반복

$$\delta_t(j) = \max_{1 \le i \le N}[\delta_{t-1}(i)a_{ij}]*b_j(o_t) \ 2 \le t \le T, 1 \le j \le N...(2.1)$$

$$\Psi_t(j) = \arg\max_{1 \le i \le N}[\delta_{t-1}(i)a_{ij}] \ 2 \le t \le T, 1 \le j \le N$$

Ⅲ. 종료

$$P^* = \max_{1 \le i \le N}[\delta_T(i)]$$

$$q_T(j) = \arg\max_{1 \le i \le N}[\delta_T(i)]$$

Ⅳ. 역추적

$$q_t^* = \Psi_{t+1}(q_{t+1}^*), \ t = T-1, \ T-2, ..., 1$$

위의 알고리즘에서 i, j는 상태 변수이며 t는 프레임 변수를 나타낸다. 연속 HMM을 기반으로 하는 본 논문을 위해 1.1의 $b_j(o_1)$는 $b_j(x_1)$으로 대치되며, 2.1의 $b_j(o_t)$는 $b_j(x_t)$로 대치된다. 위의 Viterbi 알고리즘은 계산시간을 줄이기 위해 모델의 대수(logarithm)값을 이용하여 구현될 수 있다.

4. HMM 기반 행동상태 예측

4.1 사용자 프로파일 데이터 셋

본 논문에서는 "MIT Placelab"에서 "행동 인식 시스템(Activity Recognition System)" 구축을 위해서 설치한 "State-Change Sensor"로 부터 생성된, 행동이 시작하고 종료된 시간과 센서가 설치된 장소로부터 발생한 데이터를 사용했다. 행동 라벨 샘플링은 ESM(Experience Sampling Method) 방식을 사용하여 행동에 대한 ID를 부여 하였다.

센서가 설치된 장소는 가전제품(TV, DVD, 스테레오 음향기기, 냉장고, 냉동기, 마이크로웨이브, 세탁기, 램프, 조명 스위치, 오븐, 버너, 커피머신, 스토브 등)이나 가구(캐비넷, 컨테이너, 의약품 캐비넷, 창문 등) 이다. (표 2)는 본 논문에서 사용한 원본 데이터이며, 행동 별로 가전제품이나 가구를 사용 한 시간정보와 센서가 설치된 장소 데이터가 저장되어 있다.

본 논문은 ESM을 사용하여 Labeling 된 행동을 Main/Sub action으로 분류를 한다. (표 3)은 부엌과 화장실에서 발생할 수 있는 Main action을 정의한 결과이며, 부엌에 설치된 센서에서 측정된 행동

(표 3) Main/Sub Action Dataset

Main action	Sub action	Code	Location
Preparing breakfast Preparing lunch Preparing dinner Preparing a snack	Burner	131	Kitchen
	Coffee machine	133	Kitchen
	Oven	138	Kitchen
	Toaster	139	Kitchen
	Refrigerator	137	Kitchen
	Freezer	144	Kitchen
	Microwave	143	Kitchen
	⋮		
Bathing Toileting	Sink faucet-hot	68	Kitchen
	Sink faucet-cold	88	Kitchen
	Shower faucet	93	Kitchen
	Exhaust Fan	96	Kitchen
	⋮		

데이터를 Sub action으로 분류하여, 본 논문에서 제안한 가중치 수식과 HMM에서 관측 데이터에 적용 한다.

4.2 가중치 수식화

본 논문에서는 거주지 사용자의 방과 방에서 발생된 행동 사이의 연관성을 객관적으로 수식화 하는 방법을 수식 1과 같이 제안한다.

$$w_i(SAct_i, _j) = a\frac{\sum_i count(SAct_i)}{count(SAct_k)} + (1-a)\frac{\sum_i durn(SAct_i)}{durn(MAct_k)} \quad \text{(수식1)}$$

$SAct_i$: *Action ID*, $Room_j$: *Room ID*
$MAct_k$: *Main Action ID*

a : 행동 횟수와 행동 지속 시간의 상대적인 중요도 $(0 < a < 1)$

$w_i(SAct_i, _i)$: 행동($SAct_i$)과 방($Room_j$)사이의 연관성

$count(SAct_k)$: 방에서 전체 행동 횟수($SAct_k$)의 합

$\sum_i count(SAct_i)$: 방에서 행동 횟수($SAct_i$)의 합

$durn(MAct_k)$: 방에서 행동($MAct_i$) 지속 시간

$\sum_i durn(SAct_i)$: 방에서 행동($SAct_i$) 지속 시간의 합

가중치 계산식에서 행동 횟수에 대한 수식은 구하고자 하는 Sub action 횟수를 Main Action에서 발생된 모든 Sub action의 합으로 나누어 계산하므로 행동 횟수에 관한 계산 결과는 0과 1 사이의 값을 갖는다. 그러나, 행동 발생 시간 수식에서 구하고자 하는 Sub action의 행동 발생 시간은 Main action이 종료된 시점에도 지속될 수 있으므로, 행동 발생 시간 계산 결과가 1이상의 값을 가질 수 있다.

위와 같은 행동 프로파일 데이터에 근거하여 (표 4)에서의 가중치 계산 결과는 객관적으로 가중치와 행동 발생 횟수, 행동 지속 시간 사이의 관계를 파악하기 위해서, 행동 횟수와 행동 지속 시간의 상대적인 중요도 변수 a를 고려하지 않았다.

4.3 HMM 기반 행동 상태 예측 방법

본 논문에서는 표 4의 dataset을 HMM에 적용하기 위해서 3개의 사용자의 행동 상태 모델을 제안한다.

가중치 계산과정에서 변수 a를 고려하지 않은 상태에서 가중치 최대값은 2에 근접하다. 그 이유는 행동 횟수에 관한 수식에서 구하고자 하는 Sub action 횟수를 Main action에서 발생된 모든 Sub action의 합으로 나누어 계산하므로 행동 횟수 결과는 0과 1사이의 값을 갖기 때문이다. 같은 방식으로, Sub action 지속 시간이 Main action 지속 시간 보다 작을 때, 행동 발생 시간 수식에 서 구하고자 하는 Sub action의 행동 발생 시간을 Main action이 지속된 행동 지속 시간으로 나누어 계산하면 0과 1 사이의 값을 가진다. 즉, 행동 횟수 최대값과 행동 지속시간 최대값의 합은 2에 근접한 가중치 값을 갖게 된다.

본 논문에서의 가중치 계산 결과 count에 대한 가중치와 duration에 관하여 약 0.75 값을 기준으로 상하 분포가 나타났으며, 두 합인 1.5를 경계로

(표 4) 가중치 계산 결과 Dataset

Mainaction	Date	Start time	End time	Duration(sec)	Room
Preparing breakfast	04/10/2003	AM 9:21:30	9:21:30	567	Kitchen
Subaction	Code	Duration(sec)	Count	Room	W
Toaster	131	7	2	Kitchen	0.2123
Coffee machine	119	77	1	Kitchen	0.2358
Cabinet	59	2	1	Kitchen	0.1035
Cabinet	73	43	1	Kitchen	0.1758
Light switch	105	11927	1	Kitchen	21.1353
Cabinet	55	1	1	Kitchen	0.1018
Drawer	84	3	1	Kitchen	0.1053
Dishwasher	70	5356	1	Kitchen	9.5462
Refrigerator	91	13	1	Kitchen	0.1229
Freezer	137	5892	1	Kitchen	10.4951

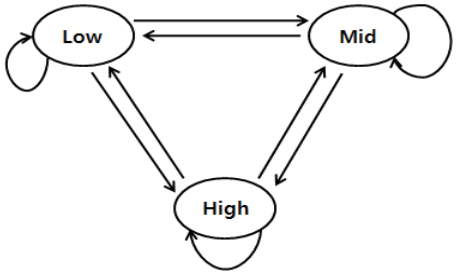
(표 5) 사용자 행동 상태 정의

Range of the WEIGHT value	User's Action State
$0 < w \leq 1.5$	Interest Low(L)
$1.5 < w \leq 2$	Interest Mid(M)
$2 < w$	Interest High(H)

한 사용자 행동 상태 정의는 다음과 같다. 두 가중치의 합이 0보다 크고 1.5보다 작거나 같은 경우 행동상태를 Low로 설정, 가중치가 1.5 보다 크고 2보다 작거나 같은 경우 행동상태를 Mid로 설정한다. 그러나 Sub action 지속 시간이 Main action 지속 시간보다 클 경우가 발생 될 수 있다. 즉, Main action이 종료된 상황에도 Sub action이 지속되었음을 의미하며 가중치는 2보다 큰 값을 나타나게 된다. 이와 같은 경우의 행동상태를 High로 설정한다.

위와 같이, 행동 상태는 Low, Mid, High 상태로 분류되며, 행동 상태는 행동 인지 센서에서 측정된 Sub action의 종료 시점에 변동된다. 위에서 정의 내린 행동 상태를 기반으로 다른 상태로 변화되는 확률은 상태 전이 확률로 구한다.

행동 상태 노드는 완전 연결 구조의 성격을 띠고 있으며 상태가 이웃 상태로 전이한 후 다시 자기 자신의 상태로 되돌아 올 수 있는 구조를 나타내므로 행동 상태 전이 모델은 Ergodic model을 사용한다.



(그림 2) 3개의 행동 상태 전이도

사용자 행동 프로파일 데이터를 HMM 인수(Parameter)에 적용하기 위한 방법으로 (표 6)의 알고리즘을 제안한다.

사용자 행동 상태를 가중치 분포 범위별로 Low/Mid/High로 구분을 하고, 행동 감지 센서에서 저장된 행동 코드(Main/Sub action code)를 이용하여 관측 테이블을 구성한다. 다음으로 측정된 행동 프로파일 데이터에서 행동 코드를 관측심벌 정의 규칙에 맞게 부여한다. 그리고 측정된 행동 프로파일 데이터로 부터 가중치를 계산하여 행동 상태 정의 규칙에 맞게 행동별로 상태를 지정한

(표 6) 사용자 행동 데이터를 HMM 인수에 적용하는 알고리즘

```
Create state definitions according to table Ⅳ
Create observation types according to table Ⅴ
Determine(time slot index, action state, observation type)
For each row
Find action(Main, Sub action)code and calculate
WEIGHT to respective action code
Assign an observation type for each time slot
following rules in Table Ⅳ
Assign an action state for each time slot
following rules given in Table Ⅴ
Determination of HMM parameters
Compute Transition Matrix A
Compute Transition Matrix B
```

다. 설정된 HMM 인수를 이용하여 초기 상태 확률 벡터, 상태 전이 확률, 관측 확률을 계산한다. 마지막으로 행동 상태를 예측하기 위해서 최적 상태 순서(optimal sequence)를 결정하는 디코더 방식의 Viterbi 알고리즘을 이용한다.

5. 실험 및 결과 분석

(표 4)의 데이터셋을 (표 6)에서 언급된 알고리즘에 적용을 하여 HMM 인수를 구한다. 그 결과의 인수값은 (표 7, 8, 9)와 같으며, 구해진 HMM 인수를 Viterbi 알고리즘에 적용을 하여 사용자의 행동 상태를 예측을 한다. (표 10)은 행동 프로파일 훈련 데이터와 예측할 행동 데이터 셋의 크기를 다르게 하여 행동상태 예측을 한 결과이다. 예측할 데이터의 크기가 15일이고 훈련 데이터의 크기를 5, 10, 15일로 설정을 했을 때, 훈련 데이터의 크기가 커질수록 정확도는 높았다.

(표 8) 초기 상태 확률 벡터($\pi = |\pi_i|$)

User's Action State	Probability
Low	0.68
Mid	0.27
High	0.05

(표 9) 상태 전이 확률($A = |a_{ij}|$)

User's Action State	Low	Mid	High
Low	0.78	0.12	0.1
Mid	0.32	0.53	0.15
High	0.42	0.37	0.21

(표 10) HMM Viterbi Decoder를 이용한 행동 상태 예측 정확도

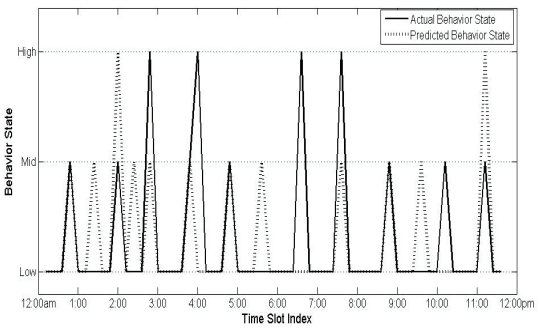
Training Set Size	Prediction Set Size	Accuracy
5 days	15 days	81.23%
10 days	15 days	86.14%
15 days	15 days	91.21%
5 days	30 days	78.67%
10 days	30 days	81.27%
15 days	30 days	87.20%

그리고 사용자의 행동 상태를 예측할 데이터의 크기를 30일로 설정하고 상태 정확도를 측정 한 결과도 같은 방식으로 훈련 데이터의 크기가 많을수록 정확도는 향상 되었지만 15일의 테스트 데이터의 크기보다 전체적으로 행동 상태를 예측하는 정확도는 낮아졌다.

사용자의 상태의 정확도를 측정하기 위하여, (그림 3)과 같이 실제 행동 상태 데이터와 예측된

(표 7) 관측 확률($B = |b_j(v_k)|$)

State	Observation Probability														
	131	133	138	139	137	144	143	68	88	93	96	79	81	57	58
Low	0.310	0.073	0.001	0.003	0.019	0.172	0.115	0.008	0.047	0.032	0.001	0.200	0.012	0.002	0.005
Mid	0.097	0.007	0.002	0.012	0.420	0.072	0.095	0.005	0.090	0.019	0.002	0.001	0.044	0.080	0.054
High	0.006	0.027	0.264	0.042	0.003	0.021	0.120	0.051	0.003	0.001	0.028	0.021	0.004	0.098	0.311



(그림 3) HMM을 이용하여 실제 행동 상태와 예상된 행동 상태

행동 상태 데이터를 시간 슬롯별로 비교 하여 행동이 발생한 같은 시간대에 행동 상태를 예측하고 정확도를 계산하였다.

6. 결론 및 향후 연구 방향

본 논문은 유비쿼터스 홈 네트워크 시스템에서 저장된 사용자 행동 프로파일 데이터에 HMM에 적용하여 사용자의 행동 상태를 예측하는 알고리즘을 제안하였다. 사용자의 행동별로 흥미를 객관적으로 측정하기 위한 방법으로 가중치를 부여했으며, 가중치를 수식화 하는 계산식을 제안 하였다. 사용자의 행동 상태를 예측하는 알고리즘으로는 디코더 방식의 Viterbi 알고리즘을 사용하였다. 실험 결과에 의하면, 훈련 데이터의 양이 많을수록 행동 상태 예측 정확도는 향상 되었지만 테스트 데이터의 양에 따라서 예측 정확도는 낮아졌다.

향후 연구로 유비쿼터스 환경에서 거주자의 행동 상태를 다른 방향으로 정의 내리는 방법과, 사용자의 상태 및 행동 패턴의 예측 정확도를 향상시키는 알고리즘을 개발할 예정이다.

참 고 문 헌

[1] 최재동, 윤원식, 민승욱, 조위덕. “유비쿼터스와 스마트 홈 서비스 기술”, 한국방송공학회, 방송공학회지, 제9권 제1호, pp.34~46, 2004

[2] Sang-Hak Lee, Tae-Choong Chung, “System architecture for context-aware home application” Software Technologies for Future Embedded and Ubiquitous Systems, 2004. Proceedings. Second IEEE Workshop, pp149-153, 2004

[3] Krumm, J, Harris, S. Meyers, B, Brumitt, B, Hale, M, Shafer, S, “Multi-camera multi-person tracking for EasyLiving” Visual Surveillance, 2000. Proceedings. Third IEEE International Workshop, pp 3- 10, 2002

[4] Morioka, K, Hashimoto, H, “Appearance based object identification for distributed vision sensors in intelligent space” Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference, pp199-204, 2005

[5] Mori, T, Hayama, N, Noguchi, H, Sato, T, “Informational support in distributed sensor environment sensing room” Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop, pp 353-358, 2005

[6] T. Mori, A. Fujii. M. Shimosaka, H. Noguchi and T. Sano, “Typical Behavior Patterns Extraction and Anomaly Detection Algorithm Based on Accumulated Home Sensor Data,” FGCN 2007, Vol 2, pp.12-18, 2007

[7] Munguia Tapia, Stephen S. Intille, Kent Larson, “Activity Recognition in the Home Using Simple and Ubiquitous Sensors”, The Second International Conference on Pervasive Computing, pp.158-175, 2004

[8] 정경용, 김종훈, 강운구, 임기욱, 이정현. “스마트 홈에서 마이닝을 이용한 행동 순차 패턴”, 한국콘텐츠학회논문지 제8권, 제9호, pp.19-26, 2008

[9] 이헌규, 노기용, 서성보, 류근호, “캘린더 패턴 기반의 시간 연관적 분류 기법”, 한국정보

과학회, 정보과학회논문지, 제32권 제6호, pp. 567~585, 2005

[10] 강형창, 양근탁, 김철수, 이윤정, 이봉규, “아이템 사용시간을 고려한 Apriori알고리즘”, 대한전기학회, 전기학회논문지, 제59권 제7호, pp.1327~1331, 2010

[11] L. R. Rabiner, “A tutorial on Hidden Markov models and selected application in speech recognition,” Proceedings of the IEEE, Vol. 77, No. 2, pp.257-286, 1989

◐ 저 자 소 개 ◑

신 동 규

1986년 서울대학교 계산통계학과(이학사)
1982년 Illinois Institute of Technology 범퓨터과학과 (공학석사)
1997년 Texas A&M University 컴퓨터과학과 (공학박사)
1998년~현재 세종대학교 컴퓨터공학과 교수
관심분야 : 상황인식 미들웨어, 웹기반 멀티미디어, 데이터베이스, 데이터마이닝
E-mail : shindk@sejong.ac.kr

신 동 일

1988년 연세대학교 전산과학과(이학사)
1993년 M.S. in Computer Scinece, Washington State University
1997년 Ph.D in Computer Scinece, University of North Texas
1998년~현재 세종대학교 컴퓨터공학과 교수
관심분야 : 상황인식 미들웨어, 무선인터넷, 게임, 지능형 에이전트, HCI
E-mail : dshin@sejong.ac.kr

황 구 연

2006년 김포전문대학 컴퓨터네트워크학과
2010년 세종사이버대학교 유비쿼터스컴퓨팅학과
2010년~현재 세종대학교 컴퓨터공학과 석사과정
관심분야 : RTLS, 유비쿼터스컴퓨팅, IPv6, 데이터마이닝
E-mail : hgy1999@gce.sejong.ac.kr

최 진 욱

2010년 강남대학교 응용수학과(공학사)
2010년~현재 세종대학교 컴퓨터공학과 석사과정
관심분야 : RTLS, 데이터마이닝, 유비쿼터스컴퓨팅
E-mail : seekpeace@gce.sejong.ac.kr