
확률 함수와 상황 전파 네트워크를 결합한 상황 데이터 생성 모델

천성표* · 김성신**

Context-data Generation Model using Probability functions and Situation Propagation Network

Seong-Pyo Cheon* · Sungshin Kim**

요 약

일상적인 상황이 주로 시간을 고려한 일정에 따라서만 발생한다고 가정하면 각 상황별 확률분포함수에 기반한 데이터 생성법은 효과적이다. 하지만, 사람들은 상황을 인식하거나 판단할 때, 이전 상황을 고려하여 현재 상황을 미루어 짐작하거나 결정하는 경우도 흔하다. 본 논문에서 제안한 상황 전파 네트워크는 상황 인식 및 결정에서 이전 상황이 현재 상황에 미치는 영향을 고려하는 것과 유사하게 상황발생 확률분포함수를 보상에 줌으로써 상황의 전개가 연속적이면서 자연스러운 흐름을 갖도록 해준다. 본 논문에서는 평범한 직장인이 집이라는 공간에서 발생시키는 상황 데이터를 생성하는 모델을 제안하였고, 모의실험을 통해 상황 전파 네트워크가 상황의 자연스러운 흐름에 얼마나 영향을 기여하는지 확인하였으며, 특히, 순차적 상황 모호성과 충돌성 지표를 새롭게 정의하여 제안한 상황 데이터 생성 모델을 평가했다. 결론적으로 상황 전파 네트워크를 결합시켰을 경우, 확률 모델만을 사용했을 때보다 순차적 상황 모호성은 6.45%, 상황 충돌성은 4.60% 감소함을 확인했다.

ABSTRACT

Probabilistic distribution functions based data generation method is very effective. Probabilistic distribution functions are defined under the assumption that daily routine contexts are mainly depended on a time-based schedule. However, daily life contexts are frequently determined by previous contexts because contexts have consistency and/or sequential flows. In order to reflect previous contexts effect, a situation propagation network is proposed in this paper. As proposed situation propagation network make parameters of related probabilistic distribution functions update, generated contexts can be more realistic and natural. Through the simulation study, proposed context-data generation model generated general outworker's data about 11 daily contexts at home. Generated data are evaluated with respect to reduction of ambiguity and confliction using newly defined indexes of ambiguity and confliction of sequential contexts. In conclusion, in case of combining situation propagation network with probabilistic distribution functions, ambiguity and confliction of data can be reduced 6.45% and 4.60% respectively.

키워드

확률기반 상황 생성 모델, 상황 전파 네트워크, 순차적 상황 모호성, 순차적 상황 충돌성

* 부산대학교 영상·IT 산학공동사업단 박사후연구원

** 부산대학교 전자전기공학과 부교수 (교신저자)

접수일자 2009. 02. 05

심사완료일자 2009. 03. 10

I. 서 론

Schilit와 Theimer는 1994년 상황인식(Context-awareness)이라는 용어를 최초로 정의했다[1]. 당시에는 상황인식을 사용자의 사용 환경이 변하면, 응용 프로그램도 새로운 환경의 변화를 반영해 줄 수 있도록 개발해야 된다고 생각하고, 이런 목적에 따라 환경의 변화를 인지하는 것이 상황인식이라고 정의했다. A. K. Day는 상황을 사용자의 감정적 상태, 주의 초점, 위치와 방향, 날짜와 시간, 물체, 그리고 사람들까지 고려한 환경이라는 개념으로 정의했다[2-3]. 이후, 상황인식과 관련한 연구의 필요성이 증가했으며, 그 영역이 점차 확대되어 독립적인 연구 분야로 자리잡기에 이르렀다. 최근에는 상황 인식을 인간과 인간 혹은 인간과 컴퓨터뿐만 아니라 물체와 물체사이의 관계까지 포함하여 상황정보를 구체화시키거나 정의하는 기술이라는 확장된 개념으로 발전시키고 있다.

한편, 유비쿼터스 컴퓨팅을 위한 인프라 구축이 상당부분 진척되면서, 구축된 환경의 활용이라는 측면에서 상황인식에 대한 관심과 필요성이 다시 한 번 부각되고 있다. 이러한 요구를 수용하기 위해서, 상황인식과 관련한 최근의 연구는 이전에 제안했던 상황인식 기술을 현실 세계에서 구현하기 위한 것과 새로운 개념으로 접근하는 상황인식에 관한 연구가 동시에 진행되고 있다. 상황인식 기술은 일반적으로 동적으로 변화하는 상황을 관리하는 기술, 상황을 추론하는 기술, 그리고 상황을 모델링하는 기술로 분류가능하다[4]. 본 논문에서는 상황 모델링 기술 가운데 상황 데이터 생성 모델을 제안한다. 임의 데이터 생성 기법을 기반으로 하는 확률 기반 상황 생성 모델은 상황 별로 확률 함수를 정의한 후, 데이터를 생성하는 방법이므로, 구조가 간단하면서도 실제 상황과 유사한 데이터가 생성되는 장점이 있다. 하지만, 이러한 임의 데이터 생성법(Random data generation method)은 상황 데이터 생성과정에서 모호성과 충돌성을 가진다는 한계가 있다. 이러한 한계를 극복하기 위해서 본 논문에서는 상황 전파 네트워크를 제안했다. 더불어, 엄격한 성능 평가를 위해서 순차적 상황 모호성과 충돌성 지표를 각각 새롭게 정의하여 성능평가를 시도하였다. 모의실험을 통해서 상황 전파 네트워크의 성능을 확인했는데, 상황 전파 네트워크와 확률 모델을 결합한 방법이 확률분포

함수만을 이용한 모델에 비하여 모호성은 6.45%, 충돌성은 4.60% 줄었다.

II. 상황 데이터 생성 방법

임의 데이터 생성 기법을 기반으로 하는 확률 기반 상황 생성 모델은 각 상황 별로 발생 가능성을 확률 분포 함수를 정의한 후, 특정 시간에 특정 상황을 임의성에 의존하여 발생시키는 방법으로 각 상황 별 확률 분포 함수만 정의하면 모델이 완성됐다고 할 정도로 간단하면서도 강력한 데이터 생성법이다. 특히, 임의 데이터 생성 기법은 대부분의 데이터 생성과정에서 광범위하게 사용되며, 직관적이기 때문에 상황 데이터 생성뿐만 아니라 표본추출이나 보간법에서도 꾸준히 활용하는 방법이다. 하지만, 확률 기반 모델이 가진 태생적 한계가 있어 지속적으로 문제가 제기되고 있다. 확률 기반 모델의 경우, 순차적 상황 변화를 반영하기 어렵기 때문에 생성된 상황 데이터를 전후 시점 또는 전후 상황의 연관성을 기준으로 판단했을 때, 모호한 경우가 많으며, 상황과 상황이 중복되거나 충돌하는 현상이 빈번히 나타난다. 이를 극복하기 위해서 제안된 대표적인 방법이 온톨로지 모델(Ontology model)이다. 온톨로지 모델은 특정 상황을 발생시키는 요건들을 미리 정의한 후, 해당 요건들이 충족될 때에 비로소 특정 상황이라고 판단하여 데이터를 생성하는 방법이다. 하지만, 인간이 실제로 상황을 짐작하거나 인식할 때는 주관적인 경우가 많으며 동일 조건에서도 심리적 또는 정황적 판단에 의해 구체적 상황을 다르게 인식하는 경우도 비일비재하기 때문에 온톨로지 모델만을 이용해서 상황을 표현하기가 힘들다. 다만, 한정된 상황들 가운데 특정 상황을 추론해야 하는 분야에서는 그 성능이 입증되고 있다. 또한, 온톨로지 모델은 고려해야 할 상황이 추가되거나 감소할 경우, 관련 요건들의 인과관계를 전체적으로 변화시켜야 하기 때문에 유연성과 확장성이 떨어지는 문제도 있다. 최근에는 온톨로지 모델의 단점을 보완하기 위해서 베이저안 네트워크(Bayesian Network)를 온톨로지 모델에 결합시켜 유연성과 추론 성능을 향상시키기 위한 시도도 이루어지고 있다[5-6]. 본 논문에서는 상황 전파 네트워크를 이용하여 확률기반 모델에서 상황들 간의 모호성과 충돌성을 극복하려는 시도를 했다.

III. 순차적 상황 흐름

인간은 상황을 인식할 때, 인식하는 순간만을 고려하지 않는다. 동일한 요인들이 갖추어진 경우에도 이전 시점에서 상황이나 이후에 해야 하는 행동에 따라서 다른 상황이라고 인식하거나 의도적으로 다른 상황으로 만들기도 한다. 본 논문에서는 이러한 실제 상황인식의 복잡성을 고려하기 위해서 상황 전과 네트워크를 제안한다. 제안한 방법은 이전 시점에서 판단한 상황이 현재 시점에 미치는 영향을 반영하기 위한 것으로 순차 행동의 사결정과정에서 사용하는 마르코프 결정 모델을 적용한 것이다.

마르코프 결정 모델은 강화학습(Reinforcement Learning)에서 일반적으로 사용하는 방법이다[7-8]. 마르코프 결정 모델은 에이전트가 각 상태에 따라 취할 수 있는 행동이 정의된 상태-행위 천이 모델(State-action Transition Model)을 정의하고, 에이전트가 취할 수 있는 여러 조합의 상태-행위 경로들 가운데 가장 목적에 부합하는 경로를 학습한다.

각 상태에서 에이전트가 선택한 행위가 다음 상태를 결정하는데 영향을 미치게 되는데, 다음 상태를 결정하는 과정은 보상 함수(Reward function)를 통해서 목적에 가장 최적화된 정책을 따르는 상태-행위로 결정하게 된다. 이것을 구체적으로 표현하면, 특정 초기 상태(Initial state) s 에서 행위(action) a 를 선택하였을 때, 그 다음 상태 s' 로의 천이과정을 모델 $T(s, a, s')$ 로 정의하고, 각 상태-행위에 따른 보상 함수 $R(s)$ 값이 가장 큰 방향으로 최적화 정책 π^* (Optimal policy)를 결정하게 한다. 본 논문에서는 상태-행위 천이 모델 $T(s, a, s')$ 을 확률 모델로 간주하여 상황 전과 네트워크로 제안하였으며, 각 상황의 변화가 다음 상황에 미치는 영향을 기대값으로 표현하였고, 식 (1)로 정의한다.

$$\begin{aligned} T(s, a, s') &= T_{ss'}^a \\ &= P\{s_{t+1} = s' | s_t = s, a_t = a\} \\ R(s, a, s') &= R_{ss'}^a \\ &= E\{r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'\} \end{aligned} \quad (1)$$

여기서, $T_{ss'}^a$ 는 t 시점의 상태가 s 이고, 그 때 에이전

트가 선택할 수 있는 행위 중에서 a 를 선택하였을 때, $t + 1$ 시점에서 에이전트가 s' 이 될 확률을 말하며, $R_{ss'}^a$ 는 위와 동일한 상태천이과정에서 에이전트가 받는 보상에 대한 기대값이다. 본 논문에서는 최적화 정책 π^* 을 구하는 것이 목적이 아니라, $T_{ss'}^a$ 와 $R_{ss'}^a$ 를 각 상태-행위에 따라 갱신시켜서, 제안한 상황 모델에 일관성과 개연성을 부여하고자 한다. 본 논문에서 제안한 모델은 마르코프 특성(Markovian property)을 따른다고 가정했기 때문에, 상태-행위 천이과정에서 생길 수 있는 시점에 대한 논란을 이전 시점으로 한정해 줌으로써, 모델의 복잡성을 줄였다[9-10].

IV. 확률 분포 함수의 정의

상황인식은 추상적인 활동이며 개인적인 견해에 의해 결정되는 경우가 흔하기 때문에, 실생활의 모든 상황을 구체적이고 객관적으로 정의하는 것은 불가능하다. 따라서, 상황인식 연구에서는 특정 기술을 사용하는 목적에 맞게 별도의 필요한 상황만을 정의한다. 상황정보를 크게 분류하면 사용자 상황, 물리·환경적 상황, 컴퓨팅 시스템의 상황, 사용자-컴퓨터 상호 작용에 대한 상황, 기타 미분류 상황으로 나눌 수 있다[11].

Anders Kofod-Petersen은 사용자의 상황을 업무 상황, 사회적인 상황, 개인적인 상황, 시간·공간적인 상황, 환경적인 상황으로 정의했다. 여기서, 개인적인 상황을 다시 생리적인 상황과 정신적인 상황으로 나누어 설명하고 있다[12]. 통용되는 상황 분류를 고려할 때, 유비쿼터스 환경에서 서비스를 위한 상황은 사용자 상황, 생리적 상황, 물리·환경적 상황, 공간 상황, 시간 상황, 활동 상황 등으로 나눌 수 있다.

본 논문은 현재 사용자가 하고 있는 행동이나 상태를 정의한 사용자 상황을 가장 중요하게 생각한다. 구체적으로, 가정이라는 공간에서 발생할 수 있는 전체 11가지 사용자 상황을 고려하였으며, 11가지 사용자 상황을 8가지의 독립적인 상황 - 외출, 청소, 학습, 식사, 요리, 설거지, 목욕/세면, 그리고 수면 - 과 이전 상황의 영향을 받는 조건부 상황 - 귀가, 휴식, 기상 - 으로 나누었다. 조건부 상황은 외출이라는 상황이 발생해야 선택될 수 있는 귀가와 특정 임무를 수행한 후 주로 선

택되는 휴식, 그리고 수면을 취해야 선택되는 기상이 있으며, 이들은 주로 이전 상황의 선택 여부에 따라 결정되는 것이 자연스럽게 때문에 이들을 조건부 상황으로 분류했다. 일반적인 직장인을 대상으로 각 상황의 발생 확률을 주중과 주말로 나누어 표 1과 표 2에 정의했다.

표 1. 주중의 상황 별 분포 함수 정의
Table. 1 Probability functions for daily contexts on a weekday

	확률함수
외출	$f_{good}(t) = \frac{1}{0.4\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-13}{4.5} \right ^{2 \times 11}}$
청소	$f_{cleaning}(t) = \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 0.3^2}}$
학습/독서	$f_{studying} = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 0.3^2}} + \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-20.5)^2}{2 \times 0.7^2}}$
식사	$f_{eating} = \frac{1}{0.6\sqrt{2\pi}} e^{-\frac{(t-8)^2}{2 \times 0.3^2}} + \frac{1}{0.5\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 0.5^2}}$
요리	$f_{cooking}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-18.5)^2}{2 \times 0.5^2}}$
설거지	$f_{disheswashing}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 0.2^2}}$
목욕/샤워	$f_{bathing} = \frac{1}{0.4\sqrt{2\pi}} e^{-\frac{(t-7.2)^2}{2 \times 0.15^2}} + \frac{1}{0.6\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 0.3^2}} + \frac{1}{0.9\sqrt{2\pi}} e^{-\frac{(t-20.5)^2}{2 \times 0.2^2}} + 0.01$
수면	$f_{sleeping}(t) = 1 - \frac{1}{0.4\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-15}{7} \right ^{2 \times 8}} - \frac{1}{3\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-3.5}{3} \right ^{2 \times 1}}$
귀가	$f_{homecoming}(t) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{(t-16.5)^2}{2 \times 0.5^2}}$
휴식	$f_{resting}(t) = \frac{1}{1.5\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-20.5}{1.35} \right ^{2 \times 2.27}}$
기상	$f_{\geq ttingup}(t) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{(t-8)^2}{2 \times 0.5^2}}$

표 2. 주말의 상황 별 분포 함수 정의
Table. 2 Probability functions for daily contexts on a weekend

	확률함수
외출	$f_{good}(t) = \frac{1}{0.8\sqrt{2\pi}} \left\{ e^{-\frac{(t-11)^2}{2 \times 1^2}} + \frac{1}{1 + \left \frac{t-16}{2.5} \right ^{2 \times 7}} \right\}$
청소	$f_{cleaning}(t) = \frac{1}{1.3\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-12}{2} \right ^{2 \times 3}} + \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-16)^2}{2 \times 1.2^2}}$
학습/독서	$f_{studying} = \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-15)^2}{2 \times 1.2^2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-19)^2}{2 \times 1^2}}$
식사	$f_{eating} = \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-10)^2}{2 \times 0.4^2}} + \frac{1}{0.6\sqrt{2\pi}} \left\{ e^{-\frac{(t-12)^2}{2 \times 0.5^2}} + e^{-\frac{(t-18)^2}{2 \times 1.6^2}} \right\}$
요리	$f_{cooking}(t) = \frac{1}{1.2\sqrt{2\pi}} e^{-\frac{(t-10)^2}{2 \times 0.6^2}} + \frac{1}{0.6\sqrt{2\pi}} e^{-\frac{(t-12)^2}{2 \times 0.5^2}} + \frac{1}{1.6\sqrt{2\pi}} e^{-\frac{(t-18)^2}{2 \times 1.6^2}}$
설거지	$f_{disheswashing}(t) = \frac{1}{\sqrt{2\pi}} \left\{ e^{-\frac{(t-12)^2}{2 \times 1.2^2}} + e^{-\frac{(t-18)^2}{2 \times 1.2^2}} \right\}$
목욕/샤워	$f_{bathing} = \frac{1}{\sqrt{2\pi}} \left\{ e^{-\frac{(t-9.2)^2}{2 \times 0.5^2}} + e^{-\frac{(t-18)^2}{2 \times 0.6^2}} \right\} + 0.1$
수면	$f_{sleeping}(t) = 1 - \frac{1}{0.45\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-11.5}{2.5} \right ^{2 \times 6}} - \frac{1}{0.45\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-18.5}{3.2} \right ^{2 \times 8}} - \frac{1}{0.3\sqrt{2\pi}} \frac{1}{1 + \left \frac{t-3.5}{3} \right ^{2 \times 1}}$
귀가	$f_{homecoming}(t) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{(t-16.5)^2}{2 \times 0.5^2}}$
휴식	$f_{resting}(t) = \frac{1}{1.2\sqrt{2\pi}} e^{-\frac{(t-12)^2}{2 \times 1.0^2}} + \frac{1}{0.8\sqrt{2\pi}} e^{-\frac{(t-15)^2}{2 \times 1.4^2}} + \frac{1}{1.2\sqrt{2\pi}} e^{-\frac{(t-20)^2}{2 \times 1.5^2}}$
기상	$f_{\geq ttingup}(t) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{(t-8)^2}{2 \times 0.5^2}}$

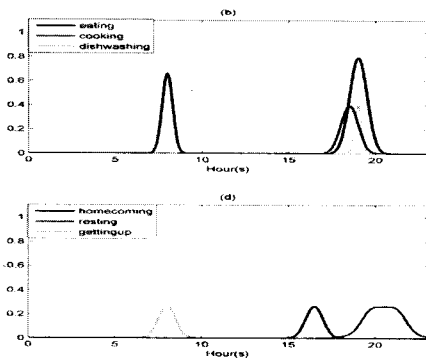


그림 1. 일과시간에 따른 주중 상황 별 확률 함수의 분포
 Fig. 1 Distributions of daily contexts probability functions on a weekday

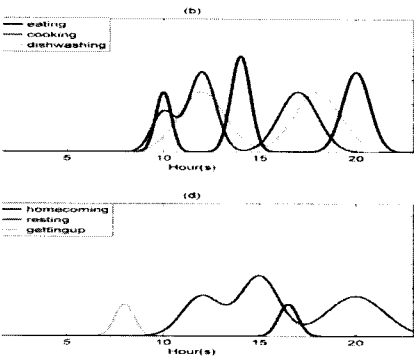


그림 2. 일과시간에 따른 주말 상황 별 확률 함수의 분포
 Fig. 2 Distributions of daily contexts probability functions on a weekend

그림 1과 그림 2는 각 상황이 발생할 가능성을 시간을 기준으로 도시했다. 그림 1의 경우 평일 직장인의 일과에 따라 오전 7시 전후로 기상하여 씻고, 간단한 식사를 마치고 직장으로 외출했다가 오후 6시를 전후로 귀가하여 씻고, 요리를 해서 저녁식사를 한 후 청소, 설거지 또는 학습/독서 시간을 가진 후, 밤 12시 전후로 잠자리에 든다. 그림 2는 주말의 경우인데, 늦잠을 자거나 낮잠도 자면서 휴식을 취하기도 하고, 집안 청소를 하는 등 대체적으로 자유로우며 불규칙적인 상황과 중복되는 경우가 많도록 모델링했다.

표 1과 표 2에서 정의한 확률함수는 특정시점에서 개별 상황이 독립적으로 발생할 확률만을 정의했다. 데이터 생성은 특정시점에서 11가지 상황이 발생할 확률의 합을 1로 정규화 한 후, 각 상황이 차지하는 비율을 계산한 후, 룰렛-휠법(Roulette-wheel method)을 이용하여 선택했다.

V. 상황 전파 네트워크

인간이 상황을 인식한 후 취하는 행동 중에는 이전 상황이나 현재 수행하고 있는 임무와의 관계에 의해 결정되는 것이 많다. 본 논문에서는 11가지 상황 가운데 귀가, 휴식, 그리고 기상을 이전 상황이나 임무의 영향을 많이 받는 조건부 상황/임무로 정의했다. 일상생활에서 ‘귀가’라는 상황은 ‘외출’의 결과로 발생하고, ‘기상’은 ‘수면’이라는 행위가 있어야 발생한다는 것을 반영한 것이다. ‘휴식’의 경우는 이전시점에 임무를 수행했을 때 그에 대한 보상으로 선택될 확률이 더욱 높아지는 것으로 가정했다.

그림 3에 제안한 확률 함수와 상황 전파 네트워크가 결합하여 상황 데이터를 생성하는 과정을 도시했다. 그림 3을 살펴보면, 사용자의 초기 확률 함수가 결정된 상태에서, 이전 시점까지의 사용자의 상황/임무가 일정 위주의 독립적 성격이었는지 혹은 이전 상황의 영향을 많이 받는 조건부 상황이었는지 여부가 상황 전파 네트워크의 입력이 된다. 상황 전파 네트워크를 통과하면서 갱신된 확률 함수가 현재시점까지의 최신 확률 함수가 되며, 이 함수를 임의 데이터 생성법을 이용하여 상황 데이터를 생성하게 되는 것이다. 그림 4는 제안한 상황 전파 네트워크이다.

그림 4에서 나타난 것처럼, 11가지 상황이 서로 연결되어 이전 상황/임무에 따라서 현재 확률함수의 변화량이 증가 또는 감소하도록 구성되어 있다. 상황 전파 네트워크는 마르코프 특성을 따르도록 설계되었으므로, 직전 시점의 영향만을 고려한다. 각 상황 노드사이의 연결가중치가 이전 시점에서의 상황이 현재 시점에 관련 상황들에 미치는 영향이다. 예를 들어, 직전 시점에서 ‘외출’이라면, 현시점에서 ‘귀가’가 선택될 확률이 10% 증가(=1.1)하고, ‘목욕/수면’을 할 확률이 20% 증가(=1.2)하게 된다.

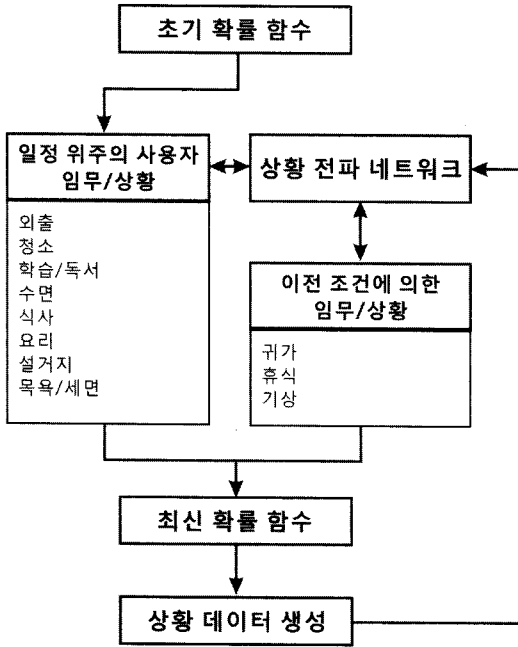


그림 3. 확률 함수와 상황 전파 네트워크의 결합을 통한 상황 데이터 생성 과정
 Fig. 3 Context data generation flowchart through combining probability functions and situation propagation network

을 가지지만, 상황 전파 네트워크에 의해서 연속으로 선택될수록 자신이 다시 선택될 확률은 20%(=0.8)씩 감소한다. 반면, '귀가'는 10%(=1.1)씩 증가한다. 그런데, 그림 1의 (d)를 보면 '귀가'의 선택 확률은 18시 이후에 비로소 값을 가지게 되고, 본 논문에서는 18시경 '귀가'한다고 가정했기 때문에, 18시를 전후하여 '외출'보다 '귀가'가 선택될 확률을 높이기 위해서 '외출'과 '귀가' 사이의 가중치를 1.1로 결정했다.

VI. 성능 평가 지표의 정의

제안한 상황 데이터 생성법은 상황 전파 네트워크를 확률 기반 모델과 결합하여 확률 기반 데이터 생성법이 가진 장점은 살리면서도 단점으로 거론되는 모호한 상황의 생성과 상황들 간의 충돌 문제를 극복하기 위한 목적으로 제안되었다. 이러한 모호성과 충돌성은 순차적으로 생성되는 상황 데이터에서만 나타나는 현상이므로, 본 논문의 성능 평가 지표는 순차적 상황 데이터가 가진 모호성과 충돌성이다. 본 논문에서 정의한 순차적 상황 모호성과 순차적 상황 충돌성을 각각 식 (2)와 식 (3)으로 나타냈다.

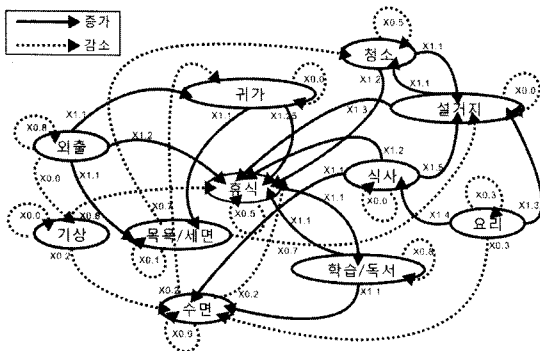


그림 4. 제안한 상황 전파 네트워크
 Fig. 4 Proposed Situation Propagation Network

만약, 연속적으로 '외출'이 선택되는 경우는 관련 노드의 현시점에서 확률은 가중치의 거듭제곱만큼 영향을 받는다. 본 논문에서 가중치의 결정은 각 상황사이의 상관관계를 고려하여 결정했다. '외출'의 경우, 그림 1의 (a)에서 9시에서 18시까지 연속으로 9시간 동일한 확률

순차적 상황 모호성 (Ambiguity)

$$= \frac{n(\{x|x(t) = x(t+1), t \in A, x \in S\})}{n(\{x|x(t), t \in A, x \in S\})} \quad (2)$$

순차적 상황 충돌성 (Confliction)

$$= \frac{n(\{x|x(t) = b \text{ and } x(t-1) = x(t+1) = a, t \in A, b \in S, a \in C\})}{n(\{x|x(t), t \in A, x \in S\})} \quad (3)$$

여기서, $n(\cdot)$ 은 조건을 만족시키는 원소의 수이며, 집합 $A = \{t|1 \leq t \leq 24, t \in \mathcal{N}\}$ 는 일과시간을 나타내며, 표준시간 단위를 1시간으로 간주했다. 집합 $S = \{\text{귀가, 휴식, 기상}\}$ 는 조건부 상황을, 집합 $C = \{\text{외출, 청소, 학습/독서, 식사, 요리, 설거지, 목욕/세면, 수면}\}$ 는 일정에 의한 상황의 집합이다. 식 (2)로 나타낸 순차적 상황 모호성은 조건부 상황/임무가 반복적으로 선택되는 횟수를 의미하며, 식 (3)의 순차적 상황 충돌성은 $t-1$ 시점과 $t+1$ 시점에 동일 상황/임무를 수

행하는데, t 시점에 조건부 상황/임무가 발생하는 경우로 정의했다.

VII. 모의실험

모의실험을 통해서, 표 1과 표 2에서 정의한 11가지 상황의 확률분포함수를 초기 확률 함수로 두고 1시간 간격으로 1년 동안의 데이터를 생성해 보았다. 데이터 분석 작업은 주중과 주말을 별도로 나누지 않고, 초기 확률 분포만을 이용하여 생성한 결과(확률 모델)와 초기 확률 분포와 상황 전과 네트워크를 결합한 경우(확률 모델 + 상황 전과 네트워크 결합 모델)를 비교했다. 그림 5에 귀가, 휴식, 그리고 기상이라는 조건부 상황에 대한 초기 확률 함수의 분포와 정규화 확률 분포를 평일과 휴일로 나누어 각각 도시했다.

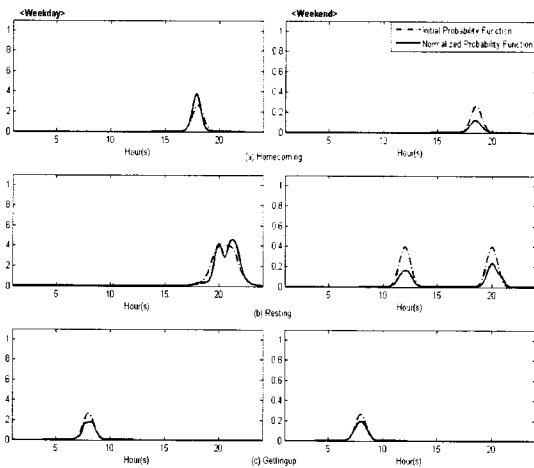


그림 5. 조건부 상황 별 초기 확률 함수의 분포 및 정규화 확률 분포

Fig. 5 Initial defined probability functions and its normalized probability distributions for conditional contexts

그림 5에서 실선으로 나타낸 것은 표 1과 표 2에서 정의한 조건부 상황에 대한 확률 함수이고, 점선으로 나타낸 것은 정규화과정을 거친 결과이다. 본 논문에서의 확률 함수 정규화는 특정 시간을 기준으로 11가지 상황들의 확률 함수값의 합을 1로 만든 후, 개별 상황이 차지하

는 확률 함수값이 차지하는 비율을 계산한 것이다. 그림 5를 보면, 평일에는 대체적으로 초기 확률 함수의 모양과 분포를 유사하게 유지하고 있으므로, 표 1에 정의한 만큼 조건부 상황이 선택될 것을 기대할 수 있다. 하지만, 휴일의 경우 분포는 비슷하지만, 함수값이 표 2에서 정의한 확률 분포값보다 높게 형성된다. 따라서, 주말에는 조건부 상황의 선택이 예상했던 것 보다 빈번히 발생할 가능성이 높다.

앞서, 식 (2)와 식 (3)에서 정의한 순차적 상황 모호성과 충돌성을 확률 모델과 확률 모델 + 상황 전과 네트워크 결합 모델을 비교하여 정리한 결과를 표 3과 표 4에 각각 나타냈다.

표 3과 표 4를 보면, ‘귀가’의 경우 확률 모델은 118회 선택했고, 결합 모델은 11회 더 많은 129회 선택했으나, 반복되는 경우와 충돌하는 경우가 각각 0회와 3회로 동일하게 나타났다. ‘휴식’의 경우 확률 모델은 399회, 결합 모델은 25회 적은 374회 선택했으나, 모호한 경우나 충돌하는 경우가 각각 37회와 18회 줄었다는 것을 확인할 수 있다. 끝으로 ‘기상’의 경우 확률 모델은 89회 선택했고, 결합 모델은 2회 적은 87회 선택했고, 모호한 경우가 4회, 충돌 상황은 11회 줄었다. 따라서, 우리는 상황 전과 네트워크를 확률 모델과 결합시킴으로써 전체적으로 모호성은 6.45%, 충돌성은 4.60% 줄일 수 있었다.

표 3. 모호성 평가 결과
Table. 3 Evaluation Result of Ambiguity

확률 모델			확률 모델 + 상황 전과 네트워크		
귀가	휴식	기상	귀가	휴식	기상
0/118	105/399	4/89	0/129	68/374	0/87
모호성=17.99%			모호성=11.54%		

표 4. 충돌성 평가 결과
Table. 4 Evaluation Result of Confliction

확률 모델			확률 모델 + 상황 전과 네트워크		
귀가	휴식	기상	귀가	휴식	기상
3/118	45/399	20/89	3/129	27/374	9/87
충돌성=11.22%			충돌성=6.62%		

VIII. 결 론

유비쿼터스 컴퓨팅이 발전함에 따라 상황인식과 관련된 기술 역시 비약적으로 발전하고 있다. 본 논문은 상황 모델링 기술 가운데 상황 데이터 생성 모델을 제안한 것이다. 본 논문에서 제안한 방법은 확률 기반 상황 생성 모델이 가진 장점은 살리면서 동시에 확률 모델이 가진 태생적 한계를 극복하기 위한 것으로, 확률 모델과 상황 전과 네트워크를 결합시킨 방법이다.

이를 통해서, 확률기반 모델이 가진 단순함과 강력함 그리고 직관적인 구현이 가능하다는 장점은 이어받으면서, 문제점으로 지적되었던 상황 모호성과 충돌성을 줄일 수 있었다. 본 논문에서는 평범한 회사원의 일상을 평일과 휴일로 나누어 확률 함수 형태로 정의한 후, 이전 상황과 현재 상황의 영향을 상황 전과 네트워크를 이용해서 반영해 주었다. 제안한 상황 데이터 생성법의 성능은 순차적 상황 모호성과 충돌성이라는 지표를 정의하여 평가하였다. 모의실험을 통해서 확률 함수만을 이용해서 생성한 데이터와 확률 함수와 상황 전과 네트워크를 결합한 모델이 생성한 데이터를 비교·평가한 결과, 상황 전과 네트워크를 이용하여 확률 모델을 보완했을 때 순차적 상황 모호성과 충돌성이 6.45%와 4.60% 줄어들었음을 확인 할 수 있었다.

참고문헌

[1] B. N. Schilit and M. M. Theimer, "Disseminating Active Map Information to Mobile Hosts," IEEE Network, vol. 8, issue 5, pp. 22-32, Sep/Oct 1994.

[2] A. K. Dey, Context-Aware Computing: The CyberDesk Project, AAAI 1998 Spring Symposium on Intelligent Environments, Technical Report SS-98-02, 1998, pp. 51-54.

[3] A. K. Dey, G. D. Abowd, and D. Salber, "A Contextbased Infrastructure for Smart Environments," Proc. Of the 1st Int' Workshop on Managing Interactions in Smart Environments (MANSE '99), pp. 114-128, Apr. 1999.

[4] J. Pascoe, N. Pyan, and D. Morse, "Issues in Developing Context-Aware Computing," LNCS, vol. 1707, pp. 208-221, 1999.

[5] 정헌만, "유비쿼터스 컴퓨팅 환경에서의 온톨로지 기반 상황 인식 미들웨어," 한국 컴퓨터정보학회 2006년 하계학술발표논문집&학회지, 제14권 1호, pp. 165-173, 2006.

[6] 정헌만, 이정현, "유비쿼터스 컴퓨터 환경에서의 상황 인식을 위한 확률 확장 온톨로지 모델," 한국 컴퓨터정보학회 논문지, 제11권 3호, pp.239-248, 2006.

[7] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction. MIT Press, MA: Cambridge, 1998.

[8] T. Mitchell, Machine Learning. McGraw Hill International Edition, 1997.

[9] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. Springer Series in Statistics, Springer, 2001.

[10] A. Doucet, N. Freitas, and N. Gordon, Sequential Monte Carlo Methods in Practice. Statistics for Engineering and Information Science, Springer, 2001.

[11] 임신영, 허재두, "상황인식 컴퓨팅 응용 기술 동향," 한국전자통신연구원, 전자통신동향분석, 제19권 5호, 2004.

[12] Anders Kofod-Petersen, A Case-Based Approach to Realising Ambient Intelligence among Agents, Thesis for the degree doctor scientiarum, 2007.

저자소개

천성표(Seong-Pyo Cheon)



2001년 : 부산대학교 전기공학과 (공학석사)
2009년 : 부산대학교 전기공학과 (공학박사)

2009년~현재 : 부산대학교 영상·IT산학공동사업단 박사후연수연구원

※ 관심분야 : 기계학습, 지능 시스템, 공정자동화 및 진단



김성신(Sungshin Kim)

1986년 : 연세대학교 전기공학과
(공학석사)

1996년 : Georgia Inst. of Technology,
전기및컴퓨터공학과
(공학박사)

1998년 ~ 현재 : 부산대학교 전자전기통신공학부
부교수

※관심분야: 지능 시스템, 데이터 마이닝, 고장예측및
진단