

확률론적 상황 예측을 위한 개선 기법들

박수영^o

광운대학교 임베디드소프트웨어공학과
psy0435@kw.ac.kr

김용혁

광운대학교 컴퓨터소프트웨어학과
yhdfly@kw.ac.kr

Some Issues for Improving Probabilistic Context Prediction

Soo-Young Park^o

Dept. of Embedded Software Engineering
Kwangwoon University, Seoul, Korea

Yong-Hyuk Kim

Dept. of Computer Science and Engineering
Kwangwoon University, Seoul, Korea

요 약

본 논문의 이전 연구[1]에서 모바일 사용자들의 지난 상황데이터(context data)를 기반으로 다음 상황 예측을 위한 확률론적 방법을 제안하였다. 이 방법은 가중치를 나타내는 멀티그래프에 속하는 새로운 그래프 구조를 통한 확률론적 이론에 근거하여 일반적인 상황을 예측한다. 본 논문에서는 이전 연구보다 더 정확한 상황 예측을 위해 목표 상황과 관련된 상황데이터만을 선별하는 특징 선택 방법과 상황데이터가 적을 경우 균등분포를 가정하여 상황의 예측가능성을 높일 수 있는 방법을 제시하였다. 이 외에도 사용자의 다양한 생활 패턴에 대한 예측방법과 이동통신사 간의 고객정보를 보안성 있게 공유하기 위한 방법을 제시한다.

1. 서 론

사용자의 요구에 따라 다양한 모바일 폰의 기능들이 개발 및 제공되고 있으며 곧 많은 사람들이 그들의 모바일 폰 없이는 살 수 없게 될 것이다[2]. 요즘 모바일 업계에서는 사용자의 과거 상황데이터를 통해 앞으로의 상황을 예측하여 사용자들에게 진보된 서비스를 제공하는 문제가 각광받고 있다. 하나의 예로 광고업자들은 효과적인 광고를 위해 특정 시간에 사용자가 어디에서 무엇을 하는 지에 대해 알고 싶어 한다[3].

이처럼 양질의 서비스를 제공하기 위해서는 사용자의 이전 상황데이터를 이용하여 다음 상황을 정확하게 예측해야 한다. 모바일 장치나 센서에 의해 수집된 모든 상황데이터를 이용할 수 있고 이 때 사용자 상황들은 통화, SMS, 사용자 활동, 위치, 날씨 정보 등으로 구성된다. Kim 등은 사용자의 기록된 이전의 상황데이터들을 분석하여 앞으로의 상황 예측을 제공하는 확률론적 방법을 제시하였다[1].

본 논문에서는 이전 연구보다 개선된 방법으로 상황 예측 가능성을 높이는 방법들을 제시한다. 이전 연구에서는 상황 예측 데이터를 조합할 때 모든 타입의 상황 데이터를 이용하였다. 그러나 본 연구에서는 보다 정확한 상황 예측을 위해 목표 상황데이터와 관련된 타입의 상황데이터만을 선택하여 예측 알고리즘에 적용하는 특징 선택법을 제시한다. 또한 사용자의 기록된 이전 상황

데이터가 적을 경우 기존의 정규분포가 아닌 균등분포를 가정하여 예측 가능성을 높이고자 한다.

다수의 모바일 이동통신사들이 중복 고객에 대한 모바일 상황 기록 데이터들을 보안성 있게 공유하여 더 나은 서비스를 제공하기 위한 방안을 제시하고 사용자의 특수한 생활 패턴에 따른 예측 방법을 고려해본다.

다음 절들에서는 본 연구의 배경이 되는 이전 연구의 확률론적 상황 예측 방법을 자세히 살펴보고 개선되어야 할 다양한 이슈들을 제시한다.

2. 이전 연구

본 논문의 이슈들을 제시하기 전에 그 이슈들의 근본으로서 Kim 등[1]이 제안했던 확률론적 상황 예측의 배경과 사용된 데이터 구조 및 예측 알고리즘에 관해서 요약 설명한다.

2.1 상황 예측의 배경

상황 예측을 위해 두 가지 타입의 상황데이터를 사용한다. 모바일 장치의 센서 정보로 수집된 감지 데이터와 사용자의 활동 정보를 나타내는 로그(log) 데이터이다. 데이터는 주기적으로 서버에게 전송되고 서버는 기록된 상황데이터를 분석하여 각 사용자의 다음 상황을 예측한다.

다양한 타입의 이전 상황데이터로부터 목표 상황 타입에 속하며 주어진 시간주기 내에 발생할 확률이 가장 큰 상황을 예측한다. 이러한 내용을 확장하여 사용자들이 어떤 주기에 따라 다른 패턴을 보인다고 가정한다 예를 들어, 사용자의 패턴은 주중과 주말에 매우 다를 수 있다.

2.2 TIMPN (Time-Inferred Multiple Pattern Network) 자료구조

상황 예측을 위해 TIMPN이라 불리는 자료구조를 이용한다. TIMPN을 방향과 가중치가 있는 멀티그래프의 유형으로서 간주할 수 있다. 각 노드는 상황 또는 상황들의 조합을 의미하며 세 가지 타입의 노드가 존재한다 먼저, k 가지의 상황 타입이 있다고 가정한다 타입1 노드는 주어진 상황 요소들을 나타내며 각 상황 타입의 타입1 노드 u 에서 머문 시간의 평균(μ_u), 편차(σ_u)와 빈도수(c_u)의 정보를 가진다 타입2 노드는 타입2 노드 u 에서의 서로 다른 상황 타입들의 순서쌍(γ_1^u, γ_2^u), 시간 차이의 평균(μ_u), 편차(σ_u)와 빈도수(c_u)의 정보를 유지한다. 빈번하게 나타나는 상황 수열로부터 앞으로의 타입2 노드를 결정하기 위해 일반적인 상황 수열($\gamma_1^u, \gamma_2^u, \dots, \gamma_{l_u}^u$), 시간 차이의 평균(μ_u), 편차(σ_u)와 빈도수(c_u)의 정보를 유지한다. 마지막으로 상황 타입 S 와 목표 상황 타입 T 사이의 다양한 에지들을 고려한다

전처리 과정에서 타입3 노드를 구성하기 위해 먼저 상황 로그로부터 자주 발생하는 작은 패턴들을 찾는다 목표 상황 타입 T 의 각각의 상황 요소 τ 에 대해, 타임 윈도우내에 τ 을 가지고 종료하는 부분열을 추출한다 이어서 각 부분열의 쌍에 대해 가장 긴 공통 부분열(longest common subsequence)을 찾는다. 그 후 Levenshtein 거리와 k -means 클러스터링을 적용하여 길이가 가능한 가장 짧은 중심 수열을 찾고 이를 타입2 노드 집합에 추가시킨다 만약 중심 수열의 길이가 1일 경우 해당 노드는 제외시킨다.

그 다음 단계로 TIMPN을 구성한다. 비어있는 TIMPN의 초기상태는 단 하나의 노드들로 구성되며 입력 상황 데이터를 하나씩 읽으면서 반복적으로 노드와 에지들을 구성한다. 각 에지를 갱신할 때, 에지에 축적된 시간 차이에 대응하는 정규분포를 검사한다 그들이 정규분포를 따르지 않는 경우, 에지를 둘이나 더 많은 에지들로 나눈다.

TIMPN의 유지보수에서는 최근의 상황데이터가 오래

된 것보다 더 중요하기 때문에 최근의 데이터에 더 가중치를 두는 에지의 노화(aging) 전략을 이용한다

2.3 예측 알고리즘

현재 시간이 t 이고 최근의 상황데이터가 주어졌을 때 주어진 시간 주기 $[t_s, t_e]$ 내에 특정 사용자의 가능한 모든 상황 요소와 그들의 확률들을 나열한다 상황을 예측하기 위해 미리 구성된 TIMPN을 사용한다. 각 목표 노드 v 의 가능성 p_v 은 다음의 세 가지 요인들에 의해 추론할 수 있다.

1) 에지 $e_{u,v,k}$ 와 관련한 각 노드 u 로부터 목표 노드 v 로의 확률 $P(u, v, k)$ 을 추정한다. 각 노드의 시간-차이 데이터와 각 노드의 수행 시간 데이터가 정규분포를 따른다고 가정하고 $P(u, v, k)$ 을 정확하게 구하기 위해서 조건부 확률과 정규합분포를 이용한다

2) 상황 예측가능성은 각 노드 u 와 목표 상황 T 사이에서 유사성 측정으로부터 파생될 수 있다 다음의 분산 측정을 이용한다 범위는 0부터 $1-1/k$ 이다.

$$p_{u,t} = \frac{\sum_k w_{u,t,k}}{\sum_{t \in T, k} w_{u,t,k}} \text{일 때, } S(u, T) = |T| \times \text{Var}_{t \in T}(p_{u,t}) \text{이다.}$$

3) 노드 u 부터 노드 v 까지 에지 $e_{u,v,k}$ 의 중요도는 에지 가중치를 기반으로 정의된다 가중치가 있는 합으로서 각 목표 노드 v 의 확률 p_v 를 정의할 수 있다

2.4 주기적인 데이터의 관리

각 기본 주기 요소 당 TIMPN을 생성하고 예측 알고리즘을 적용하기 전에 유사한 패턴을 보이는 그룹의 TIMPN들을 하나로 결합해야 한다. 어느 두 개의 TIMPN $G_A = (N_A, E_A)$ 와 $G_B = (N_B, E_B)$ 사이의 유사성 $s(G_A, G_B)$ 은 에지들의 정규분포들 사이의 유사성 $s(e^A, e^B)$ 들의 정규화된 합계로부터 얻을 수 있다

정규분포 $N(\mu_1, \sigma_1)$ 과 $N(\mu_2, \sigma_2)$ 를 따르는 에지 e_1 과 e_2 에 대한 유사성을 구하기 위해 다음의 세 가지 경우를 고려하였다. 첫 번째 경우는 에지 e_1 과 e_2 가 똑같은 분포를 따르는 경우이다. 이 경우 $s(e_1, e_2) = 1$ 이다. 두 번째 경우는 e_1 과 e_2 의 확률 밀도 함수가 하나의 교차점을 갖는 경우이다. 일반성을 잃지 않고 $\mu_1 < \mu_2$ 을 가정하면 교차점이 $(\mu_1 + \mu_2)/2$ 일 때, $s(e_1, e_2) = 2 \cdot P\left(z \leq \frac{\mu_1 - \mu_2}{2\sigma}\right)$ 이다. 세 번째 경우는 e_1 과 e_2 의 확률 밀도 함수가 두

개의 교차점을 갖는 경우이다. 이 경우 일반성을 잃지 않고 $\sigma_1 < \sigma_2$ 를 가정한다.

$$t_1 = \frac{b - \sqrt{b^2 - ac}}{a}, t_2 = \frac{b + \sqrt{b^2 - ac}}{a} \text{ 이고}$$

$a = \sigma_2^2 - \sigma_1^2, b = \mu_1\sigma_2^2 - \mu_2\sigma_1^2, c = \mu_1^2\sigma_2^2 - \mu_2^2\sigma_1^2 + 2\sigma_1^2\sigma_2^2 \ln \frac{\sigma_1}{\sigma_2}$ 일 때,

$$s(e_1, e_2) = P\left(z \leq \frac{t_1 - \mu_1}{\sigma_1}\right) + P\left(\frac{t_1 - \mu_2}{\sigma_2} \leq z \leq \frac{t_2 - \mu_2}{\sigma_2}\right) + P\left(z \geq \frac{t_2 - \mu_1}{\sigma_1}\right)$$

이 된다.

3. 개선 이슈

3.1 상황데이터의 특징 선택 (feature selection)

특징 선택[4,5]은 예측의 성능을 높이기 위해 의미 있는 특징들만을 최적으로 추출하는 방법이다. 이는 인공지능과 데이터마이닝 분야에서 발생하는 다양한 예측 문제에 널리 쓰이는 기법이다.

본 논문에서는 상황 타입을 예측하기 위하여 이전의 상황데이터들을 고려할 때 목표 상황과 관련된 상황데이터를 선별하여 TIMPN 알고리즘에 적용한다. 목표 상황과 관련 없는 상황데이터 타입들은 예측가능성을 낮추고 그릇된 결과를 도출하는 원인이 될 수 있으므로 특징 선택을 통해 이러한 요인들을 제거한다. 예를 들어 사용자의 '위치' 상황데이터를 예측하기 위해 관련성이 없다고 알려진 '음악' 상황데이터를 이용할 경우 오히려 노이즈(noise)로 작용하여 정확하지 않은 예측 결과가 나올 수 있다. 이러한 문제를 예방하는 특징 선택의 방법론은 다음과 같다.

상황 예측가능성에 대한 가중치를 구하기 위해 각 노드 x 와 목표 상황 집합 Y 사이의 유사성 $S(x, Y)$ 측정을 이용한다. 목표 상황 Y 의 노드들의 개수가 k 일 때의 변화량을 측정하여 다음의 결과를 얻을 수 있다

$$0 \leq S(x, Y) = k \times \text{Var}_y(p_{x,y}) < 1.$$

목표 상황 Y 의 k 개 노드가 동일한 확률을 가질 경우 $S(x, Y)$ 는 0의 최소값을 가지고, k 개의 노드 중 한 노드만이 선택된다면 $S(x, Y)$ 는 $1 - 1/k$ 의 최대값을 가지게 된다. 특징 선택을 위한 평가 함수 $J(X)$ 가 주어지고 예측될 각 타입1 상황에 대한 특징 선택이 이루어진다. 그림 1은 특징 선택을 빠르게 수행하기 위한 의사코드(pseudo-code)를 보여준다.

평가 함수 $J(X)$ 를 측정하기 위한 방법은 다음과 같다. 먼저 목표 상황데이터가 포함된 훈련 데이터(training data)가 주어진다. 상황데이터의 타입들을 나타내는 X 를

```
// We have already known the similarity S(x, Y)
Sort features Xs in non-increasing order of E_X(S(X, Y));
F_0 = ∅;
for(i = 1; i ≤ K; i++) { // K : # of features
    F_i = F_{i-1} ∪ {i^{th} feature};
    f_i = J(F_i);
}
return arg max_{F_i} f_i;
```

그림 1. 빠른 특징 선택을 위한 의사코드

특징 부분집합으로 이용할 때 TIMPN으로 변형 후 시간 주기 내에서 가장 발생 확률이 높은 목표 상황데이터 타입을 선별하여 예측 정도를 가늠한다.

3.2 빈도가 적은 상황데이터의 향상된 처리 기법

과거의 기록 데이터의 양이 적을 경우 빈도가 낮은 데이터에 정규분포 가정을 하기가 어렵기 때문에 확률 분포 값이 항상 일정한 경우인 균등분포(uniform distribution)를 가정하여 예측 가능성을 높일 수 있다.

이전 연구(2.3절)에서 예측 알고리즘을 위한 정규분포의 경우를 다루었고 이 절에서는 개선 방법으로서 균등분포를 적용한 균등분포의 경우와 균등분포와 정규분포의 합을 고려한다.

첫 번째 경우로 균등분포를 구해보겠다. U_1 은 $U[l_1, u_1]$ 의 범위 안에 있고, U_2 은 $U[l_2, u_2]$ 의 범위 안에 있다고 가정한다. U_1 의 범위를 $-l_1$ 만큼 이동시킨 $U_1^* = U_1 - l_1$ 은 $U[0, u_1 - l_1]$ 의 범위 안에 있고, U_2 의 범위를 $-l_2$ 만큼 이동시킨 $U_2^* = U_2 - l_2$ 은 $U[0, u_2 - l_2]$ 의 범위 안에 있다. 일반성을 잃지 않고 $u_1 - l_1 \leq u_2 - l_2$ 라 가정하면 이 둘의 합은 $U_1^* + U_2^* = U_1 + U_2 - (l_1 - l_2)$ 이고 $U_1 + U_2 = U_1^* + U_2^* + (l_1 + l_2)$ 라고 할 수 있다. 누적분포함수는 다음과 같이 나타낼 수 있다.

$$P(U_1^* + U_2^* \leq t) = \int_0^t P(U_1^* = x)P(U_2^* \leq t - x)dx.$$

누적분포함수에 t 에 관한 미분을 취하면 다음의 확률밀도함수를 얻을 수 있다(그림 2).

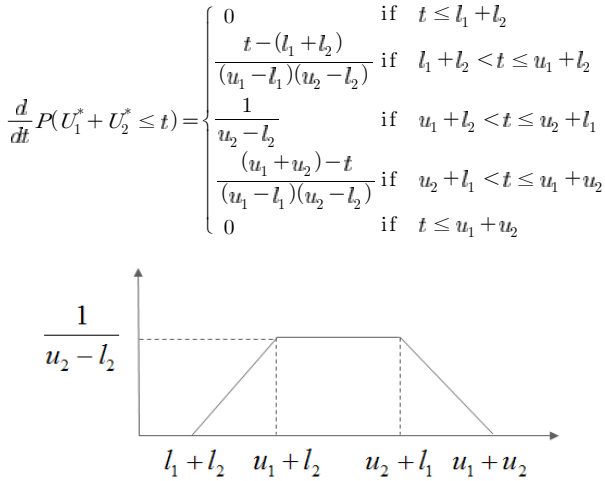


그림 2. 균등분포의 확률밀도함수 그래프

두 번째 경우로 균등분포와 정규분포의 합을 알아보겠다. 균등분포 U 는 $U[l, u]$ 의 범위에 있고 정규분포 N 은 $N(\mu, \sigma^2)$ 의 범위에 있다고 가정한다 폐쇄형(closed form)으로서의 확률밀도함수는 얻을 수 없다 다행스럽게도 다음의 누적분포함수를 이용하기만 하면 된다

$$P(U + N \leq t) = \int_{-\infty}^{\infty} P(N = x)P(U \leq t - x)dx.$$

$$P(U + N \leq t) = P\left(z \leq \frac{t - u - \mu}{\sigma}\right) \cdot \frac{u - t + \mu}{u - l} + p\left(z \leq \frac{t - l - \mu}{\sigma}\right) \cdot \frac{t - l - \mu}{u - l} + \frac{\sigma^2}{u - l} \cdot (f(t - l) - f(t - u)).$$

$$\text{단, } f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

이전 연구(2.4절)에서 주기적인 데이터의 관리를 위해 두 정규분포를 따르는 예지들의 병합 과정을 고려하였다. 본 논문에서는 두 균등분포 사이의 유사성과 균등분포와 정규분포 사이의 유사성을 고려한다

첫 번째 경우, 두 균등분포 사이의 유사성 $s(e_1, e_2)$ 는 다음과 같다.

$$s(e_1, e_2) = \min(h_1, h_2) \times (\min(u_1, u_2) - \max(l_1, l_2)).$$

두 번째 경우, 균등분포와 정규분포 사이의 유사성 $s(e_1, e_2)$ 는 $t = \mu \pm \sigma \sqrt{2 \ln \frac{1}{\sqrt{2\pi} \cdot \sigma h}}$ 일 때, $h = 1/(u - l)$ 의 범위에 따라 그림 3의 네 가지의 경우로 나누어 볼 수 있다.

3.3 통신사 간의 보안성 있는 정보 공유

사용자가 현재 이용하고 있는 이동통신사의 서비스를

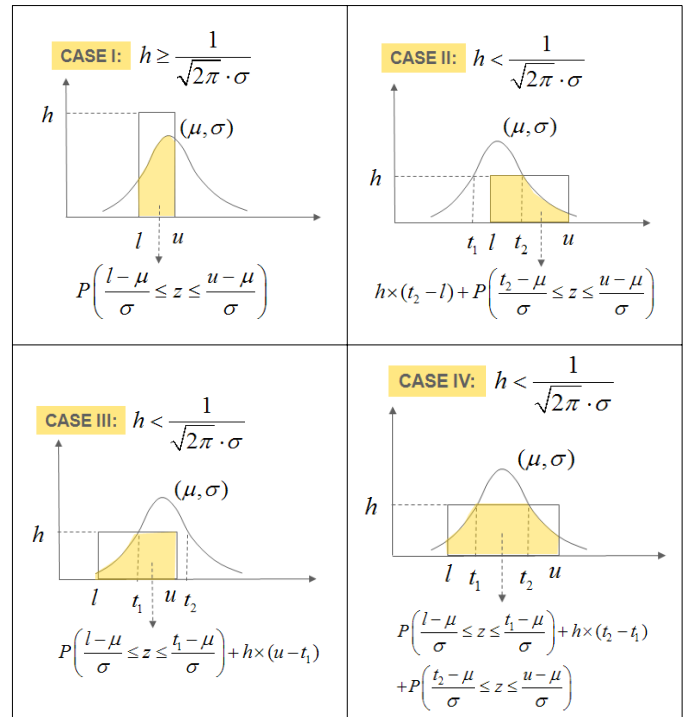


그림 3. 균등분포와 정규분포 사이의 유사성

다른 이동통신사로 바꿀 경우, 새로운 이동통신사는 사용자의 이전 모바일 기록 정보를 통해 진보된 서비스를 제공하고자 할 것이다 또한 동일 사용자가 다수 이동통신사의 서비스를 이용할 경우 해당 이동통신사들 간에 사용자의 모바일 기록 정보를 공유함으로써 더 나은 서비스를 제공할 수 있다

이처럼 이동통신사들 간에 동시 가입된 고객 또는 새로 가입된 고객의 모바일 기록정보들을 암호화하여 공유할 수 있다면 고객 정보의 유출의 위험을 방지하면서 본 논문에서 다른 사용자의 상황데이터 분석을 통한 모바일 서비스의 질을 향상시킬 수 있을 것이다 다음에 소개하는 특허[6]를 이용하여 위의 보안유지 방법을 실현하기 위한 방법을 제안한다

이동통신사는 고객들의 명단과 모바일 사용기록을 합법적으로 항상 저장 및 유지하고 있다고 가정한다 다수의 이동통신사와 그들의 정보를 처리 및 전송할 주관사가 존재한다 고객이 모바일 이동통신사를 A 업체에서 B 업체로 바꾸었을 경우, 해당 고객의 암호화된 정보가 A 업체에서 주관사로 보내진다 B 업체 또한 암호화된 새 고객의 정보를 암호화하여 주관사로 보내고 주관사는 전송된 고객들의 암호화된 정보들을 복호화하지 않은 상태에서 중복 고객을 추출하여 A 업체에 저장되어 있는 구 고객의 상황데이터 정보를 B 업체로 보낸다(그림 4).

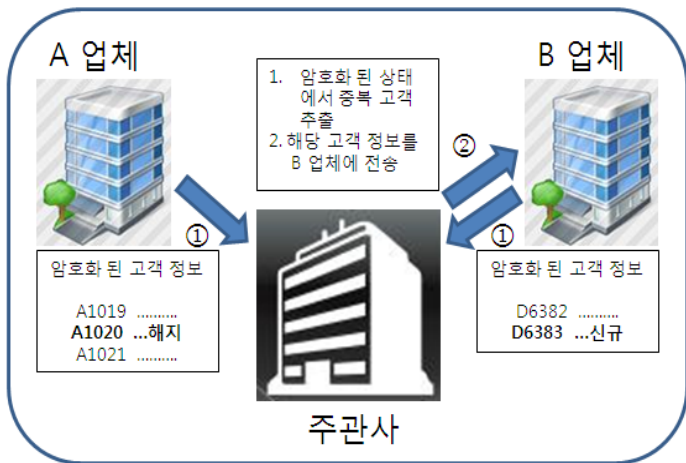


그림 4. 보안성 있는 정보 공유의 예

정보의 암호화 방식은 개인을 식별할 수 있는 고유정보만 암호화하고 일반적인 상황데이터는 암호화하지 않는다. 예를 들어 고객의 이름과 주민번호 휴대폰 번호만 암호화하고 모바일 정보 사용기록은 그대로 보내는 방식이다.

이러한 방법으로 사용자의 이전 데이터 기록정보들을 이동통신사와 주관사 모두 복호화하지 않은 상태에서 중복 고객을 판별하고 정보를 전송하여 보안성을 유지하며 공유한다. 특허[6]에 따르면 위의 공유시스템을 주관사 없이 다수 참여기업들만으로 구성할 수도 있다

3.4 특수한 생활 패턴을 보이는 사용자의 처리

예측한 상황데이터 적중률의 임계값(threshold)을 설정해놓고 임계값 보다 낮은 적중률이 나타날 경우 이전의 분포들을 저장하고 새로운 TIMPN을 설계한다.

1) 비주기적인 생활 패턴을 보이는 사용자의 경우

목표 상황과 예측된 상황의 적중률을 갱신시키면서 임계값 이하로 낮아질 경우 생활 패턴이 변화했다고 판단한다. 따라서 기존의 생활패턴에 대한 예측방법을 저장해놓고 새로운 TIMPN을 설계한다. 이를 확장하여 다음의 경우를 가정할 수 있다

2) 다중적인 생활 패턴을 보이는 사용자의 경우

감지 방법은 앞의 경우와 같으며 새로운 TIMPN과 기존의 TIMPN이 예측한 상황과 목표 상황과의 적중률을 주시하여 또 다른 주기적인 변화가 있는지 감지한다 주기적으로 패턴 그래프 구조가 반복될 경우 반복되는 작은 패턴들을 조합하여 전체적인 새로운 주기적인 TIMPN을 생성한다.

이처럼 주기적인 생활 패턴이 아닌 특수한 생활 패턴

을 보이는 사용자에게 대한 주기적인 상황 데이터를 처리할 때, 목표 상황데이터의 적중률의 임계값을 설정하여 보다 신속하고 정확하게 TIMPN을 설계할 수 있으며 비주기적인 생활 패턴을 큰 주기로 확장하여 TIMPN을 적용할 수 있다.

4. 결론

본 연구에서는 모바일 사용자들의 과거의 기록된 상황 데이터를 기반으로 현재 상황을 예측하는 이전의 TIMPN 접근법에 새로운 사항들을 추가하여 예측 가능성을 높일 수 있는 방안들을 제시하였다

첫 번째 이슈는 목표 상황 데이터를 추정할 때 관련된 상황 타입들만을 선별하여 예측 가능성을 높이는 특징 선택이다. 이를 위해 각 노드들과 목표 상황과의 유사성을 측정하였고 평가 함수를 이용하였다

두 번째 이슈에서는 과거 상황데이터의 기록이 적은 경우, 예측 가능성을 높이기 위해 정규분포 대신 균등분포를 가정하여 상황을 예측하였다 이전 연구[1]와 비교하여 변경된 부분을 수식으로 새롭게 유도하였다

세 번째 이슈는 이동통신사 간에 보유고객의 상황데이터를 보안성 있게 공유하기 위한 방법을 제시하였고 마지막으로 특수한 생활 패턴을 보이는 사용자의 상황 예측방법에 대하여 고려하였다

이러한 이슈들을 이용하여 이전 연구의 예측 알고리즘의 정확도를 향상시키고 모바일 사용자에게 진보된 서비스를 제공할 수 있을 것으로 기대한다

참고 문헌

- [1] Y.-H. Kim, W. Kim, K. Min, and Y. Yoon, Probabilistic Context Prediction Using Time-Inferred Multiple Pattern Networks, In *Proceedings of the Annual ACM Symposium on Applied Computing*, pp. 94-98, 2010.
- [2] L. Palen, M. Salzman, and E. Youngs. Going Wireless: Behavior & Practice of New Mobile Phone Users. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 201-210, 2000.
- [3] C. Ververidis and G. C. Polyzos. Mobile Marketing Using Location Based Services. In *Proceedings of the First International Conference on Mobile-Business*, 2002.

- [4] I.-S. Oh, J.-S. Lee, and B.-R. Moon, Hybrid Genetic Algorithms for Feature Selection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 11, pp. 1424-1437, 2004.
- [5] Y.-H. Kim, S.-Y. Lee, and B.-R. Moon, A Genetic Approach for Gene Selection on Microarray Expression Data, In *Proceedings of the Genetic and Evolutionary Computation Conference - Lecture Notes in Computer Science 3102*, Vol. 1, pp. 346-355, 2004.
- [6] 문병로, 보안성을 가지는 고객정보 공유 시스템 및 그 방법, KR-A-1020030015519, 02.25, 2003.