

센서 레지스트리 시스템을 위한 3-간선 패턴 기반 경로 예측 알고리즘

이석훈*, 정동원**†, 정현준*, 백두권*†

*고려대학교 컴퓨터전파통신공학과

**군산대학교 통계컴퓨터과학과

e-mail : leha82@korea.ac.kr, djeong@kunsan.ac.kr, {darkspen, baikdk}@korea.ac.kr

Three-Edge Pattern based Path Prediction Algorithm for Sensor Registry System

Sukhoon Lee*, Dongwon Jeong**†, Hyunjun Jung*, Doo-Kwon Baik*†

*Dept. of Computer and Radio Communications, Korea University

**Dept. of Statistics and Computer Science, Kunsan National University

요 약

센서 레지스트리 시스템(SRS, Sensor Registry System)은 이기종 센서 네트워크에서 끊임 없는 의미 처리를 위하여 사용자에게 센서 정보를 제공한다. 불안정한 네트워크 상황에서의 원활한 서비스 제공을 위하여 빠른 근거리 사용자 이동 경로 예측 알고리즘(FCR, Fast and Close-Range Prediction) 기반의 SRS가 연구되었다. 이 연구는 경로 예측 기반의 SRS에서 이용되는 FCR 알고리즘이 지니는 한계를 극복하기 위하여 3-간선 패턴(TEP, Three-Edge Pattern) 기반의 경로 예측 알고리즘을 제안한다. TEP 알고리즘은 경로를 그래프로 표현할 때 사용자의 위치를 기준으로 이전 간선, 현재 간선, 다음 간선으로 패턴화 하여 학습하고, 이 패턴을 기반으로 하는 사용자의 이동 경로를 예측한다. 또한 실험 및 비교 평가에서, TEP 알고리즘이 FCR 알고리즘에 비해 높은 정확성을 지님을 보인다.

1. 서론

최근 사물인터넷(IoT) 패러다임이 많은 주목을 받으면서 통신, 교통, 헬스케어, 스마트 홈 등의 다양한 영역에서 센서 기술 및 센서 데이터 처리기술이 활발히 연구되고 있다[1]. SWE (Sensor Web Enablement)[2], SSNO (Semantic Sensor Network Ontology)[3]는 센서 데이터의 구조 및 의미를 웹에서 표현하고 제공해 주기 위한 표준들이다. 또한 이기종 센서 네트워크에서 의미 해석 및 처리를 위하여 센서의 의미 정보를 정의하고 명세하기 위한 연구들이 진행되었다[4].

센서 레지스트리 시스템(SRS, Sensor Registry System)은 이러한 센서의 의미 정보들을 사용자에게 제공함으로써 이기종 센서 네트워크 환경에서 끊임 없는 의미 해석 및 처리를 가능하게 한다[5]. SRS는 센서 의미 정보를 관리하고 공유하기 위한 시스템으로 실시간 서비스 제공을 고려하여 개발되었다. 하지만 SRS는 불필요한 센서 데이터 수신, 끊임 없는 의미 해석 및 처리 보장, 센서 정보 요청 시 과부하, 이동 단말 자

원의 최적화 문제를 지니고 있어, 이로 인한 SRS의 확장 방안 역시 연구되고 있다[6].

한편, SRS의 한계를 극복하기 위한 방안으로 경로 예측 기반의 확장된 SRS가 개발되었다[7]. 경로 예측 기반 SRS는 기존 SRS에서 불안정한 네트워크 상황에서 발생하는 의미 해석 및 처리의 끊임 현상을 개선한다. 이러한 접근법은 SRS에서 사용자의 이동 경로를 예측하고 예측된 경로에 있는 센서 정보들을 미리 적재(Preloading)함으로써 사용자가 불안정한 네트워크 상태의 지역으로 이동하더라도 안정적인 의미 해석 및 처리를 가능하게 한다.

경로 예측 기반 SRS의 구현을 위해서는 사용자와 서버 간 즉시적인 처리, 동적 센서 정보 관리 및 확장, 높은 재사용성 등과 같이 기존 SRS가 지니는 장점들을 유지하면서 경로 예측 기법 적용에 따른 장점을 얻을 수 있어야 한다. 이러한 조건을 만족시키기 위해 [7]에서는 빠른 근거리 사용자 이동 경로 예측(FCR, Fast and Close-Range Prediction) 알고리즘[8]을 이용한다. 하지만 FCR 알고리즘은 단순하고 휴리스틱한 방법으로 인해 예측의 정확도가 낮다는 한계를 지닌다.

이 논문은 경로 예측 기반의 SRS를 위하여 기존의 FCR 알고리즘이 지니는 한계를 극복하기 위하여 3-간선 패턴(TEP, Three-Edge Pattern) 기반 경로 예측 알고리즘을 제안한다. 3-간선 패턴이란 사용자가 이동하

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2014R1A1A2058992). 또한 이 논문은 2015년 정부(미래창조과학부)의 재원으로 한국연구재단-차세대정보·컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(NRF-2012M3CA7033346).

† 공동교신저자 (co-corresponding authors)

는 도로망을 하나의 그래프로 보고, 이 그래프 내에서의 연속된 3개의 간선을 패턴화 한 것을 의미한다. TEP 알고리즘은 사용자가 실제 이동한 경로를 3-간선 패턴으로 만들어 학습하고, 학습된 패턴 및 가중치를 이용하여 사용자의 근거리 경로를 예측한다. 또한 실험 및 결과를 통하여 제안 알고리즘의 장점을 기술한다.

2. 관련 연구

앞서 언급했듯 경로 예측 기반의 SRS 기술은 기존의 SRS이 지니는 장점들을 취하면서 경로 예측 연산이 수행되어야 한다. [9]는 이러한 관점에서 기존의 경로 예측 기법들을 분석하며, SRS에 적용시키기 위한 경로 예측 기법들의 장단점들을 기술한다. 기존의 경로 예측 기법들은 대부분이 전체 경로를 예측하거나 개인화된 경로를 예측한다. 하지만 이러한 경로 예측 기법들은 사용자의 과거 데이터가 존재하지 않을 경우 적용하지 못하거나, 실시간으로 예측 결과를 모니터링 하며 잘못된 경로를 수정하지 못한다. 또한 동적인 경로 예측 및 수정이 이루어지더라도 전체 경로를 예측하기 위해 많은 리소스를 차지한다.

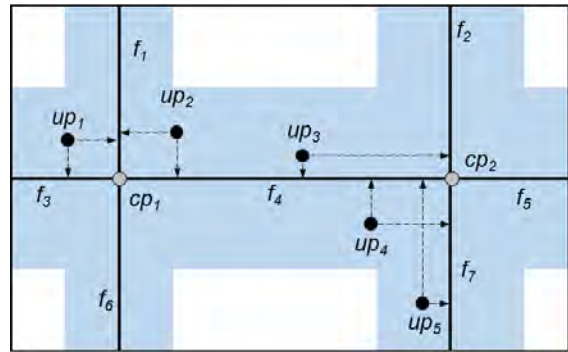
FCR 알고리즘[8]은 이러한 문제 해결들을 해결하고자 고안한 단순하지만 빠르고 즉시적인 처리가 가능한 알고리즘이다. FCR 알고리즘은 도로들의 교차점을 기준으로 도로를 경로 단편으로 나누고 경로 단편 단위로 예측이 진행된다. 또한, GPS를 이용하여 실시간으로 사용자의 위치를 읽어 근거리 경로를 빠르게 예측한다. 하지만 FCR 알고리즘은 사용자의 현재 위치만으로 가중치에 영향을 주는 다른 요인이 없어 낮은 정확성을 보인다.

[10]은 시간 요소를 이용하여 FCR의 낮은 정확성을 개선한다. 하루를 6개의 시간 구간으로 나누어 각각의 경로 단편에 대한 가중치를 개별적으로 측정한다. 이는 시간대에 따라 사용자들이 하루 동안 이동하는 패턴이 달라지는 것을 고려한 방법이다. 예를 들면, 사용자들이 출근시간, 점심시간, 퇴근시간 등에 따라 달라지는 이동 패턴을 예측에 적용하는 것과 같다. [10]은 대학 및 인근 대학로를 한정하여 FCR 알고리즘 및 시간 요소 분석 알고리즘을 구체적으로 평가한다. 하지만 이는 기존의 FCR이 지니는 예측에 시간 요소를 추가한 연구로서, 알고리즘 개선을 위해 이동 경로를 대상으로 하지 않는다.

3. FCR 알고리즘과 경로 단편의 그래프 표현

FCR 알고리즘은 도로를 교차로 단위로 단편화 시킨 경로 단편(Path Fragment)을 이용한다. (그림 1)은 FCR 알고리즘에서 사용하는 사용자의 위치와 경로를 표현한다. up_i 는 사용자의 순차적인 위치를 의미한다. 파란색 음영은 실제 사용자가 걸어 다닐 수 있는 도로를 표현하며, cp_i 는 각 도로가 교차되는 지점, f_i 는 두 cp_i 가 연결되는 경로 단편을 나타낸다.

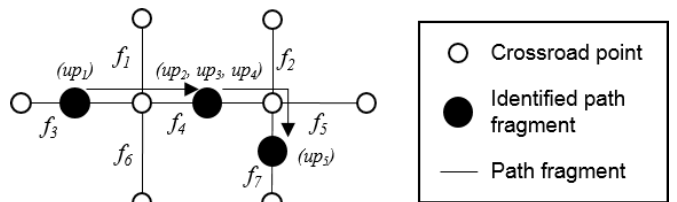
FCR 알고리즘은 사용자의 위치를 이용하여 현재 사용자가 존재하는 경로 단편을 식별하고, 식별된 경



(그림 1) FCR 알고리즘의 사용자 위치 및 경로 표현

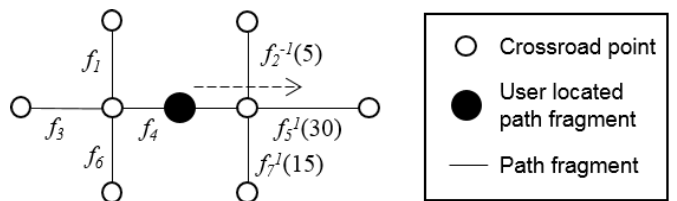
로 단편과 연결된 경로 단편을 예측한다. 이 때, FCR 알고리즘은 탐욕 전략 알고리즘(Greedy Algorithm)에 기반을 둔다. 즉 현재 위치한 경로 단편과 연결된 경로 단편들이 지니는 가중치 중에서 가장 큰 가중치를 선택하는 방식으로 다음 경로 단편을 예측한다. 예를 들면, 현재 사용자가 경로 단편 f_4 를 이동하고 있을 경우, FCR 알고리즘은 f_2, f_5, f_7 에서 가장 높은 가중치를 지니는 경로 단편을 이동할 경로로 예측한다. 각 경로 단편이 가지는 가중치는 모든 사용자들이 경로 단편을 이동한 빈도수로 측정된다.

(그림 2)는 (그림 1)에서 표현하고 있는 사용자 위치 및 경로 단편을 그래프로 표현한 그림이다. 각 원점들은 도로의 교차점을 의미하고 간선들은 교차점을 연결하는 도로의 경로 단편을 의미한다. 검은 점은 사용자 위치에 의해 식별된 경로 단편으로, 사용자가 현재 어느 경로 단편에 위치하는지를 보인다.(그림 1)에서 사용자는 사용자 위치 $up_1 \rightarrow up_2 \rightarrow up_3 \rightarrow up_4 \rightarrow up_5$ 를 순차적으로 이동하는데, 이는 (그림 2)에서 경로 단편 $f_3 \rightarrow f_4 \rightarrow f_7$ 를 이동하는 것으로 표현된다.



(그림 2) 사용자 위치 및 경로의 그래프 표현

(그림 3)은 FCR 알고리즘의 경로 예측 방식을 그래프로 표현한다. 경로 단편은 $f_i^d(w)$ 와 같이 나타내며, i 는 경로 단편의 번호를, d 는 경로 단편의 사용자 이동 예측 방향을, w 는 가중치를 의미한다. 만약 사용자가 경로단편 f_4 에 위치할 때, FCR 알고리즘은 사용자가 이동할 다음 경로 단편으로 f_5 를 예측한다.

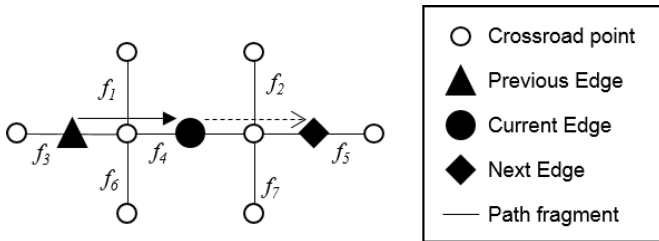


(그림 3) FCR 알고리즘의 경로 예측 방식

4. 3-간선 패턴 기반의 경로 예측 알고리즘

이 논문은 FCR 알고리즘의 단순성으로 인해 발생하는 낮은 예측 정확성 문제를 개선하기 위하여 3-간선 패턴을 학습하고 이를 이용하여 경로를 예측하는 알고리즘을 제안한다. 제안 알고리즘(TEP 알고리즘)은 FCR이 지나는 장점을 취하고 단점을 개선하는 것을 목표로 한다.

TEP 알고리즘은 경로 단편을 예측하고자 할 때, 사용자가 현재 위치한 간선과 현재 간선으로 이동하기 전에 어느 간선으로부터 이동해 왔고, 다음은 어느 간선으로 이동할 것인지를 하나의 패턴으로 인식한다. (그림 4)는 이러한 3-간선 패턴의 예를 보인다. 사용자는 f_3 을 거쳐 현재 f_4 에 위치해 있으며 이후 f_5 로 이동함을 의미한다. TEP 알고리즘은 이러한 3-간선 패턴마다 가중치를 부여한다. 사용자의 이동 경로를 분석하여 3-간선 패턴이 얼마나 빈번히 발생하는지를 학습한다. 경로 예측을 위하여 이전 간선과 현재 간선을 기준으로 가장 빈번히 발생하는 3-간선 패턴을 선택하는 방식으로 예측을 진행한다.



(그림 4) 3-간선 패턴의 예

<표 1>은 3-간선 패턴과 가중치를 보인다. 예를 들면, 사용자는 이전 간선(Previous edge) f_3 을 거쳐 현재 간선(Current edge) f_4 에 위치한다고 할 때, 다음으로 이동할 예측 간선은 f_2, f_5, f_7 중 하나이다. 3-간선 패턴으로 변환하여 보면 $\{f_3-f_4-f_2, f_3-f_4-f_5, f_3-f_4-f_7\}$ 중 가중치가 가장 큰 패턴을 선택하여 다음 간선을 예측하는데, <표 1>의 4, 5, 6번 패턴 중에서 5번 패턴의 가중치가 가장 크므로 TEP 알고리즘은 5번 패턴의 다음 간선(Next edge), 즉 f_5 를 사용자가 다음으로 이동할 예측 경로 단편으로 선정한다.

<표 1> 3-간선 패턴과 가중치

Pattern no.	Previous edge	Current edge	Next edge	Weight
1	f_1	f_4	f_2	20
2	f_1	f_4	f_5	5
3	f_1	f_4	f_7	5
4	f_3	f_4	f_2	10
5	f_3	f_4	f_5	30
6	f_3	f_4	f_7	5
7	f_6	f_4	f_2	5
8	f_6	f_4	f_5	10
9	f_6	f_4	f_7	15
...

<표 2>는 사용자의 이동 경로로부터 3-간선 패턴의 가중치 측정 알고리즘이다. 사용자가 이동한 경로들의 집합을 Path라고 하며 UP는 하나의 경로에서 사용자가 순차적으로 지나온 위치 데이터의 집합이다. 하나의 사용자 위치로부터 현재 위치한 간선(경로 단편)을 식별하여 idf로 할당하며 3-간선 패턴의 이전 간선, 현재 간선, 다음 간선을 각각 $e1, e2, e3$ 로 할당한다.

<표 2> TEP 가중치 측정 알고리즘

```

Algorithm measuring_TEP_weight (Path [])
Initialize (w[][][])
for each Path: p do
    UP [] ← get_user_points (p)
    e1 ← null
    e2 ← null
    for each UP: up do
        idf ← get_identified_path_fragment (up)
        e3 ← get_path_fragment_number (idf)
        if e2 ≠ e3 then
            if e1 ≠ null and e2 ≠ null then
                w[e1][e2][e3] ← w[e1][e2][e3] + 1
            endif
            e1 ← e2
            e2 ← e3
        endif
    endfor
endfor
return w[][][]
    
```

<표 3>은 3-간선 패턴 기반 경로 예측 알고리즘으로 사용자의 이전 간선과 현재 간선을 이용하여 현재 간선에 인접한 간선과 패턴을 구성한다. 구성된 패턴들 중에서 가중치가 가장 큰 패턴을 예측 경로 단편으로 선정한다.

<표 3> TEP 경로 예측 알고리즘

```

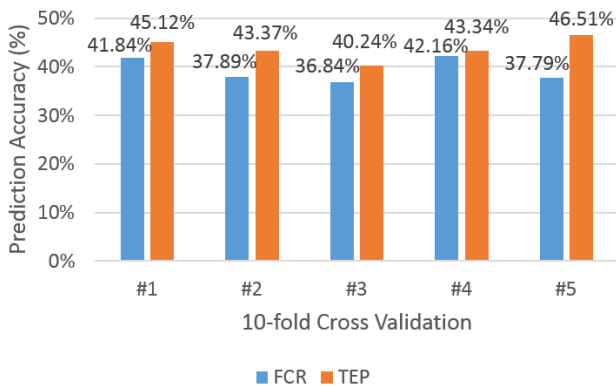
Algorithm path_prediction (up, e1, w[][][])
idf ← get_identified_path_fragment (up)
e2 ← get_path_fragment_number (idf)
CF[] ← get_connected_fragments (idf)
pf ← null
maxweight ← 0
for each CF: cf do
    e3 ← get_path_fragment_number (cf)
    if maxweight < w[e1][e2][e3] then
        maxweight ← w[e1][e2][e3]
        pf ← cf
    endif
endfor
return pf
    
```

TEP 알고리즘은 현재 간선과 이전 간선만 지니고 있으면 되므로 실시간 경로 예측이 가능하며, 단순히 현재 위치한 경로 단편과 그 방향성만으로 가중치를 구분하는 FCR 알고리즘과도 비교하였을 때 오버헤드에서 큰 차이를 보이지 않는다.

5. 비교 평가

이 장은 실험을 통하여 FCR 알고리즘과 TEP 알고리즘의 정확성과 처리 속도를 측정하고 비교 평가한다. 실험을 위하여 임의의 대학 캠퍼스와 인근 대학로의 도로를 단편화 시키고, 경로 단편을 미리 정의한다. GPS를 이용하여 5명의 실험자에게 총 10일 간 사용자의 위치를 추적하여 총 117개의 경로와 5871개의 사용자 위치 데이터를 수집하였다. 이 데이터는 실외에서만 측정하고 사용자가 한 지점에서 다른 지점으로 이동할 경우에만 수집되도록 정제하였다.

비교 대상은 FCR 알고리즘과 TEP 알고리즘을 선정하였으며, 평가 요소로는 예측의 정확성과 경로 단편의 가중치 학습 시간 및 예측 시간으로 정의한다. 정확성의 측정을 위하여 사용자 이동 경로 위주로 10-fold 교차 검증 기법을 이용한다.



(그림 5) FCR 알고리즘과 TEP 알고리즘의 정확도

(그림 5)는 10-fold 교차 검증 기법을 5번 이용하여 측정된 평균 경로 예측의 정확도를 보인다. 랜덤으로 학습 경로와 테스트 경로를 9:1의 비율로 선정하고, FCR 알고리즘과 TEP 알고리즘에 동일한 경로를 적용하여 실험하였다. 총 5번의 실험 동안 FCR 알고리즘의 경우 최대 51.85%, 최소 21.05%의 정확도 차이를 보이며, TEP의 경우 최대 61.76%, 최소 21.05%의 정확도 차이를 보여 둘 다 안정되지 않은 정확도를 보인다. 하지만 전체적으로는 TEP 알고리즘이 FCR 알고리즘보다는 높은 정확도를 보임을 알 수 있다.

<표 4>는 5번의 실험에 따른 평균적인 실험 결과를 보인다. FCR 알고리즘은 평균 39.30%의 정확도를 보이며, TEP 알고리즘은 평균 43.72%의 정확도를 보여 평균적으로 4.42%p의 정확도 개선이 일어났음을 알 수 있다. 또한 평균 학습 시간 (Average Training time) 및 평균 테스트 시간 (Average testing time)은 평균 -16 ms의 차이를 보여 두 알고리즘의 가중치 측정 및 예측 시간에 큰 차이가 없음을 알 수 있다.

<표 4> 비교 평가 결과

Algorithm	FCR	TEP	Difference
Average accuracy	39.30%	43.72%	4.42%p
Average training time	7,918 ms	7,902 ms	-16 ms
Average testing time	817 ms	817 ms	-

6. 결론

이 논문은 경로 예측 기반의 센서 레지스트리 시스템에서 FCR 알고리즘의 정확도를 개선한 3-간선 패턴 기반의 경로 예측 알고리즘을 제안하였다.

실험 결과 TEP 알고리즘은 FCR 알고리즘에 비해 처리 속도적인 측면에서 큰 오버헤드의 증가 없이 일정 부분 정확성이 개선되었음을 확인할 수 있었다. 하지만 각각의 최소, 최대 정확도의 차이가 약 30%p, 40%p 정도였으며 이는 두 알고리즘 모두 예측 결과가 안정적이지 않음을 의미한다. 그러나 향후 충분한 실험자들이 보다 많은 실험 데이터를 이용하여 예측 연산을 수행할 경우 보다 더 안정적인 결과가 얻어질 것으로 예상된다.

향후 연구로서, 제2장에서 언급한 시간 요소를 이용하여 TEP 알고리즘에 적용하기 위한 연구가 요구된다.

참고문헌

- [1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54, pp. 2787-2805, 2010.
- [2] C. Reed, M. Botts, G. Percivall, and J. Davidson, "OGC Sensor Web Enablement: Overview And High Level Architecture," Open Geospatial Consortium, 2013.
- [3] M. Compton et al., "The SSN ontology of the W3C semantic sensor network incubator group," Web Semantics: Science, Services and Agents on the World Wide Web, vol. 17, pp. 25-32, 2012.
- [4] M. Compton, C. Henson, L. Lefort, H. Neuhaus, and A. Sheth, "A survey of the semantic specification of sensors," Proc. of 2nd International Semantic Sensor Networks Workshop, International Workshop on Semantic Sensor Networks 2009, 2009.
- [5] D. Jeong and J. Ji, "A Registration and Management System for Consistently Interpreting Semantics of Sensor Information in Heterogeneous Sensor Network Environments," Journal of KIISE : Databases, vol. 38, no. 5, pp. 289-302, 2011. (in Korean)
- [6] D. Jeong, S. Lee, H. Jung, and D.-K. Baik, "Extending the Sensor Registry System for a Heterogeneous IoT Environment," Proc. of 2015 Winter Joint Conference, The Korean Association of Computer Education and Korea Society of Computer Information, vol. 19, no.1, pp.185-188, Jun, 2015. (in Korean)
- [7] D. Jeong and M. Doo, "A Path Prediction-Based Sensor Registry System for Stable Use of Sensor Information," Journal of KIISE, vol.42, no.2, pp.255-263, Feb. 2015. (in Korean)
- [8] D. Jeong, "Fast and Close-Range Prediction Algorithm of User Paths," Journal of KIISE: Computing Practices and Letters, vol. 20, no. 3, pp. 153-158, 2014. (in Korean)
- [9] H. Park and D. Jeong, "A Survey on Moving Path Prediction Methods for Improving the Performance of SRS," Proc. of the KIISE Korea Computer Congress 2013, pp.442-444, Jun. 2013. (in Korean)
- [10] M. Doo, "Time Pattern Analysis based Path Prediction Algorithm for Enhancing the Precision of FCR," Master Thesis, Kunsan National University, Korea, 2015.