

최적 증가: IoST에서 Terrestrial-CubeSat 간 핸드오버를 위한 기계학습 기반 단말기 이동성 예측 알고리즘

오준석, Anh Tien Tran, 하태윤, 조성래

중앙대학교 컴퓨터공학과

{jsoh, attran, tyha}@uclab.re.kr, srcho@cau.ac.kr

Optimal Augmentation: Machine Learning-based Terminal Mobility Prediction Algorithm for Handover between Terrestrial-CubeSat in IoST

Junsuk Oh, Anh Tien Tran, Taeyun Ha, Sungrae Cho

Department of Computer Science and Engineering, Chung-Ang Univ.

요약

CubeSat 기반의 IoST(Internet of Space Things)에서 Terrestrial-CubeSat 간 효율적인 핸드오버를 위해 단말기의 이동성을 예측하는 것은 중요한 문제로 여겨지고 있으나 이러한 네트워크 변화에 대응하는 NDP(Neighbor Discovery Protocol) 등 기존의 방식은 이를 효율적으로 지원하지 못하는 상황이다. 본 논문에서는 선형 회귀 알고리즘을 적용한 기계학습과 데이터의 최적 증가를 바탕으로 단말기의 이동성을 예측하는 알고리즘을 제안한다.

I. 서론

IoT(Internet of Things)는 무선 통신을 바탕으로 각종 사물을 연결하여 글로벌 유비쿼터스 플랫폼을 실현하기 위한 차세대 무선 시스템의 주요한 부분이다. 그러나 대기, 태양, 사막, 극지, 우주 등에서는 통신 인프라가 제공되지 않음에 따라 이러한 IoT의 잠재력을 실현할 수 없는 상황이다. 그림 1의 IoST는 CubeSat을 기반으로 이러한 IoT의 글로벌 연결을 실현하기 위한 유비쿼터스 사이버 물리적 시스템으로 정의된다 [1].

CubeSat은 부피 1ℓ(10cm×10cm×10cm), 질량 1.33kg을 초과하지 않는 초소형 인공위성을 의미한다 [2]. CubeSat의 주요 특징 중 하나인 동시성은 해당된 지역만을 대상으로 임무를 수행하는 대형 인공위성과는 다르게 동시에 지구 전체 또는 우주 전체 방향을 대상으로 특정 임무를 수행할 수 있음을 의미한다. 그러나 CubeSat이 지구 전체를 대상으로 임무를 수행할 경우 단말기의 이동성을 바탕으로 Terrestrial-CubeSat 간 링크를 재설정하기 위한 핸드오버 기술이 필수적으로 요구된다.

지상망을 거치지 않고 CubeSat과 직접 통신을 이루는 단말기의 경우 이동 경로가 항상 유동적임에 따라 이러한 즉각적인 핸드오버 처리를 위해서는 단말기의 이동성을 예측하는 기법이 필수적으로 요구된다. 이에 따라 본 논문에서는 기계학습과 최적 증가 기법을 바탕으로 단말기의 이동성을 예측하여 Terrestrial-CubeSat 간 핸드오버를 지원하고자 한다.

본 논문의 나머지는 다음과 같이 구성된다. 본문의 II-1은 단말기의 이동성으로 인해 발생하는 문제에 대해서 설명한다. 다음으로 II-2는 기계학습 기반의 단말기 이동성 예측 알고리즘을 제안한다. II-3은 데이터의 최적 증가 기법에 대해 제안하고, II-4에서 제안한 알고리즘과 기법에 대한 실험을 진행한다. 마지막으로 결론에서는 본 논문을 마무리한다.

II. 본론

II-1. DTN에서 단말기 이동성에 의한 NDP의 문제

NDP(Neighbor Discovery Protocol)는 네트워크에서 결정한 주기마다 Discovery 메시지를 전송하여 네트워크에 연결된 단말기가 이동하는 등

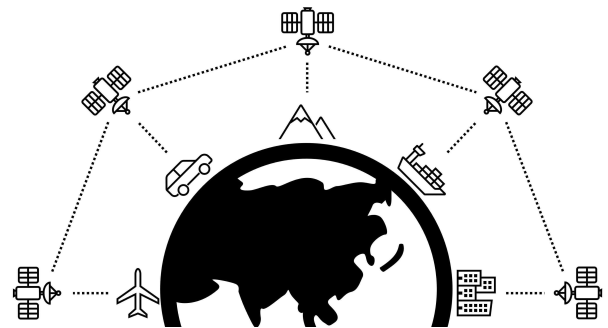


그림 1. CubeSat을 기반으로 하는 IoST

의 네트워크 변화가 발생할 경우 이에 대응하는 프로토콜을 의미한다 [3]. 그러나 IoST 등 DTN(Delay Tolerant Network)에서 NDP를 사용할 경우 단말기의 이동과 같은 즉각적인 핸드오버를 처리하기 어렵다는 문제점이 발생하게 된다. 예를 들어 특정 단말기가 네트워크에 연결된 상태에서 이동함에 따라 연결이 단절된 경우 네트워크가 NDP를 바탕으로 네트워크 변화가 발생하였음을 확인하기 전까지는 해당 단말을 복구할 수 없게 된다. 결론적으로 네트워크 변화가 업데이트될 때까지 해당 단말기의 데이터는 전달되지 않으며, 단말기가 연속적으로 이동하는 경우 네트워크가 NDP를 바탕으로 네트워크 변화를 확인하더라도 업데이트 시점의 위치와 실제 위치가 달라 해당 문제가 지속적으로 발생하게 된다.

II-2. 기계학습 기반 단말기 이동성 예측 알고리즘

단말기의 이동성을 분석하고 예측하는 것이 목적임에 따라 이를 회귀 문제로 정의하여 선형 회귀 알고리즘을 적용하고자 한다. 선형 회귀 알고리즘의 독립 변수 x 와 종속 변수 y 는 다음과 같이 정의할 수 있다.

$$x = \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}, y = \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$$

이는 단말기가 시점 x_i 에서 위치 y_i 에 존재하였음을 의미한다. x 와 y 의 집합은 선형 회귀 알고리즘에 따른 이동성 분석을 위한 훈련 데이터 집합으로 사용한다. 그러나 훈련 데이터 집합은 비선형성을 포함하고 있으므로 다음과 같은 선형 회귀 모델에 적합하지 않다.

$$y_i = \beta_1 x_1 + \dots + \beta_p x_p + \epsilon_i, \quad i = 1, \dots, n$$

이에 따라 x 에 대해 다항 기저함수를 적용하여 x 에 대해서는 비선형적이고 β 에 대해서는 선형적인 모델을 다음과 같이 정의할 수 있다.

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p$$

매개변수 β 는 MSE(Mean Squared Error)를 최소화하기 위한 가중치 값으로 정의할 수 있다. 다음으로 Noise에 대한 과소적합과 과대적합을 방지해야 한다. 이를 위해 K-fold Cross Validation을 적용하여 기저함수의 차수를 결정하고자 한다. 데이터 수가 적은 경우 정확도를 향상시킬 수 있는 K-fold Cross Validation은 훈련 데이터 집합을 1개의 Test Fold와 K-1개의 Train Fold로 분할하여 진행하는 교차 검증으로 정의된다. 기저함수의 차수가 증가할수록 함수의 차수가 증가하게 되므로 MSE는 지속적으로 감소하지만 일정 기준 이상으로 기저함수의 차수가 증가할 경우 오히려 MSE가 증가할 수 있다. 이에 따라 MSE가 최소인 기저함수의 차수를 독립 변수 x 에 대해 가장 적합한 모델로 결정할 수 있다.

II-3. 높은 정확도를 위한 데이터의 최적 증가 기법

데이터 집합인 과거 이동 경로는 불연속적인 데이터를 내포하고 있을 수 있다. 이에 따라 불연속을 연속으로 변환함과 동시에 기저함수의 차수를 최소화하는 최적 증가 기법을 다음과 같이 제안하고자 한다.

$$y_j = a_k y_i + (1 - a_k) y_{i+1}, \quad a_k < 1$$

위에서 j 는 i 번째와 $i+1$ 번째 사이에 존재하지 않는 연속적인 데이터를 표현하기 위해 추가하는 데이터이므로 i , $i+1$ 와 같지 않다. 추가하는 데이터는 i 번째 데이터와 $i+1$ 번째 데이터에 대해 가중치를 부여하여 생성하고 가중치의 합은 1을 초과하지 않는다.

$$a_{k+1} = a_k + r, \quad r < 1$$

위에서 다음 가중치 a_{k+1} 는 비율 r 을 더하여 생성하므로, 이는 데이터를 얼마나 증가시킬지 결정하는 변수로 정의할 수 있다. 결론적으로 기저함수의 차수를 최소화하며 제안한 알고리즘의 정확도를 높이는 r 을 구하는 것이 최적 증가의 목표라고 정의할 수 있다.

II-4. 단말기 이동성 예측 알고리즘과 최적 증가 기법의 실험

데이터 집합은 버스의 과거 이동 경로를 사용한다. 그림 2는 최적 기법을 적용하지 않은 경우의 기저함수의 차수에 따른 정확도와 MSE를 나타내며, 기저함수의 차수가 22인 순간부터 수렴하는 것을 확인할 수 있다. 결과적으로 최적 증가 변수 r 을 구하기 위한 기준은 기저함수의 차수인 22로 정의할 수 있다. 그림 3은 최적 증가 변수 r 에 따른 테스트 데이터 집합에 대한 정확도와 기저함수의 차수를 나타낸다. 최적 증가를 적용한 경우 r 이 0.1일 때 가장 높은 정확도를 보이는 것을 확인할 수 있으며, 최적 증가를 적용하지 않은 경우보다 기저함수의 차수가 낮아지는 것을 확인할 수 있다. 결론적으로 기준보다 기저함수의 차수를 최소화하면서 제안한 알고리즘의 정확도를 높이는 최적 증가 변수 r 을 구할 수 있음을 확인할 수 있다.

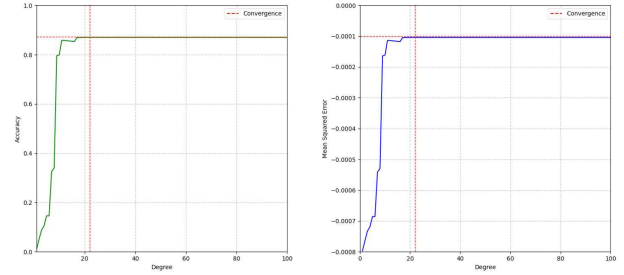


그림 2. 최적 증가를 적용하지 않은 경우의 정확도와 MSE

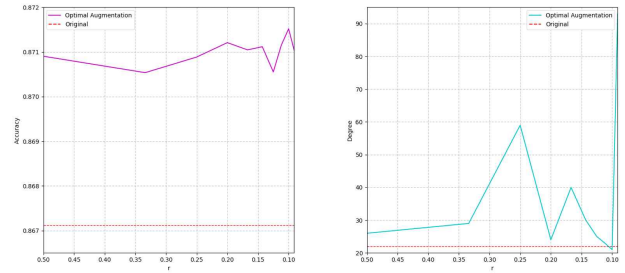


그림 3. 최적 증가를 적용한 경우의 정확도와 기저함수의 차수

III. 결론

CubeSat 기반 IoST에서 네트워크 변화에 대응하는 기존의 방식은 단말기의 이동성에 따라 발생하는 Terrestrial-CubeSat 간 핸드오버를 효과적으로 지원할 수 없다. 이에 따라 본 논문에서는 효율적으로 핸드오버를 지원하기 위해 선형 회귀 알고리즘을 적용한 기계학습 기반의 단말기 이동성 예측 알고리즘을 제안하고, 제안한 알고리즘의 기저함수의 차수를 최소화함과 동시에 정확도를 높이기 위한 데이터 최적 증가 기법을 제안한다. 실험에서 제안한 알고리즘은 높은 정확도를 제공하였으며, 최적 증가 변수를 구할 경우 제안한 기법은 제안한 알고리즘의 기저함수의 차수를 최소화하는 동시에 더 높은 정확도를 제공할 수 있음을 보인다.

ACKNOWLEDGMENT

본 연구는 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 [NRF-2019R1A2C1090447]

참 고 문 헌

- [1] I. F. Akyildiz and A. Kak, "The Internet of Space Things/CubeSats," in *IEEE Network*, vol. 33, no. 5, pp. 212-218, Sept.-Oct. 2019.
- [2] N. Saeed, A. Elzanaty, H. Almorad, H. Dahrouj, T. Y. Al-Naffouri and M. -S. Alouini, "CubeSat Communications: Recent Advances and Future Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1839-1862, thirdquarter 2020.
- [3] A. S. A. Mohamed Sid Ahmed, R. Hassan and N. E. Othman, "IPv6 Neighbor Discovery Protocol Specifications, Threats and Countermeasures: A Survey," in *IEEE Access*, vol. 5, pp. 18187-18210, 2017.