

기계학습 기반의 VNF 최적 배치 모델에 관한 연구

박수현^o, 홍지범, 유재형, 홍원기

포항공과대학교 컴퓨터공학과

{sh.park11, hosewq, jhyoo78, jwkhong}@postech.ac.kr

A Study on Machine Learning-based Optimal VNF Placement

Suhyun Park^o, Jibum Hong, Jae-Hyoung Yoo, James Won-Ki Hong

Department of Computer Science and Engineering, POSTECH

요약

NFV (Network Function Virtualization) 의 환경에서는 범용 서버에 소프트웨어로 구현되는 VNF (Virtualized Network Function)을 네트워크 토폴로지 상 적절한 위치의 노드에 배치하고 네트워크 상황에 따라 동적으로 관리함으로써, 다양한 트래픽 상황에 대해 신속하고 유연하게 대응할 수 있다. 하지만 비용과 서비스 품질 등을 고려하여 최적의 VNF 배치를 결정하고 적용하는 것은 복잡하고 어려운 문제이다. 특히, 결정된 배치를 실제 NFV 환경에 적용하는 데는 처리 시간이 소요되므로 배치가 적용되는 미래 시점의 상황을 예측하여 미리 결정하는 것이 필요하다. 본 논문에서는 MEC (Multi-access Edge Computing) 토폴로지에서 서비스 요청을 임의로 생성하여 ILP (Integer Linear Programming) 모델을 통해 시뮬레이션한 결과를 훈련 데이터로 사용하는 기계학습 모델을 도출한 후 테스트베드에서 성능을 검증하는 연구와 최종 모델을 적용한 VNF 배치 자동화 시스템을 제안한다. 이를 통해 VNF의 배치를 효율적으로 관리할 수 있다.

I. 서론

인터넷에서는 점점 더 멀티미디어를 사용하는 다양한 형태의 high-throughput 서비스 요구가 증가하고 있다. 전통적인 네트워크 환경에서는 서비스 사업자(service providers, SPs)가 서비스 처리에 필요한 기능을 제공하는 미들 박스 (middle boxes) 장비를 네트워크 내부의 적절한 위치에 추가로 설치하고 이에 맞추어, 트래픽 흐름을 제어하는 등의 작업을 수행해왔다. NFV 기술은 이러한 고가의 미들박스들을 범용 서버에 가상화하여 구현한 VNF 들로 대체하고 실시간으로 변화하는 서비스 요구에 유연하게 처리하도록 하는 기술이다 [1].

NFV 기술은 동적으로 변화하는 네트워크 서비스 요청 (service request)에 대해 유연하게 대처할 수 있는 장점이 있지만, 이를 위해 비용과 서비스 품질을 고려하여 VNF 를 동적으로 배치하고 그에 따라 SFC (Service Function Chain)를 구성하는 문제는 네트워크 관리자가 수작업으로 수행하기 어려울 만큼 규모가 크고 복잡해진다. 이는 네트워크 토폴로지, 서버의 자원 할당, 네트워크 대역폭 할당, VNF 타입별 카탈로그와 각 시점의 서비스 요청 정보를 바탕으로 비용과 서비스 품질 등의 관점에서 최적의 해를 찾는 문제로 접근할 수 있다. 하지만 결정된 VNF 배치를 적용하는 데 처리 시간 (VNF deployment time)이 소요되기 때문에 VNF 배치 결정이 이루어지는 시점의 정보로 도출한 최적의 배치가 실제 적용되는 시점에는 최적이지 아닐 수 있다는 문제가 있다 [2].

본 논문에서는 MEC 토폴로지에서 VNF 최적 배치를 결정하는 기계 학습 모델을 제안한다. 제안하는 방법은 서버의 자원 할당, 네트워크 대역폭 할당, VNF 타입별

카탈로그와 각 시점의 서비스 요청 정보를 바탕으로 최적의 VNF 배치를 산출하여 레이블링한 데이터를 생성한 후, 기계학습 기법을 통해 현재 시점의 정보를 입력으로 미래 시점의 최적 VNF 배치를 출력하는 모델을 학습시키는 것이다. 제안하는 모델의 성능이 검증되면, VNF 배치 결정을 자동으로 수행하는 시스템에 적용할 수 있다.

II. 관련 연구

M. F. Bari [3]는 VNF 배치와 SFC 구성을 결정하는 문제를 서비스 수준 협약 (Service Level Agreement, SLA) 기준을 맞추는 범위에서 네트워크 운용 비용과 이용률을 최적화 시키는 문제로 정의하고, 이 문제에 대한 ILP 를 수학적 최적화 소프트웨어 패키지인 CPLEX 에서 구현하였다. 이때 네트워크 토폴로지, 서버의 자원 할당, 네트워크 대역폭 할당, VNF 타입별 카탈로그와 각 시점의 서비스 요청 정보를 활용하였다.

X. Zhang [2]은 VNF 배치에 대한 기존 연구들이 이미 도착한 서비스 요청에 대응하는 방식으로 접근하고 있음을 지적하며, 결정된 배치를 적용하는 처리 시간을 고려하여 서비스 요청을 도착 전에 미리 예측할 필요가 있다고 주장하였다. 이를 토대로 X. Zhang 은 VNF 배치를 선제적으로 계획하는 방법을 제시하였다. 마찬가지로 B. Li [1]도 미래의 VNF-SC (virtual network function service chain) 요청을 정확히 예측하기 위해 LSTM 기반의 딥러닝 (deep learning) 모델을 설계하였다.

S. Lange [4]는 시뮬레이션으로 구한 서비스 요청 데이터와 연구 [3]의 ILP 솔루션을 기반으로 미래 시점의 VNF 타입 별 가상 서버의 소요량을 예측하였다. 이 연구는 시뮬레이션 환경에서 서비스 요청 데이터를

생성하고, 이를 기반으로 생성한 데이터를 각 시점에서의 트래픽 데이터로 변환하여 ILP 솔루션의 입력 값으로 사용하였다. 또한 이 ILP 솔루션의 결과를 기반으로 각 시점의 VNF 타입 별 가상 서버 수량 증감에 대한 결정을 레이블링하였다. 이 데이터를 바탕으로 과거부터 현재 시점까지의 통계 정보를 입력 값으로 받아 예측 시간대의 VNF 타입 별 수량의 결정을 증가/유지/감소 결과 값으로 출력하는 기계학습 기반의 분류 모델을 제안하였다. 제안하는 모델은 60 초 이후의 시점에서 최적의 VNF 타입 별 수량 증감에 대해 ILP 결과 대비 75~80%의 정확도를 달성하였다.

III. VNF 최적 배치 예측 모델

본 논문에서 제안하는 VNF 배치 예측의 형태는 그림 1 과 같다. 현재 시점 (t)에서 p 시간 이후 시점에 대한 최적의 VNF 배치를 예측하기 위해 두 가지 방법을 고려할 수 있다. 첫 번째 방법은, a)와 같이 미래 시점의 유효한 서비스 요청을 예측한 후, 이를 바탕으로 VNF 배치를 최적으로 계산한다. 두 번째 방법은, b)와 같이 새로운 서비스 요청이 도착하는 각 시점에 대해 서버 노드를 행으로 하고 VNF 타입을 열로 하는 VNF 최적 배치 행렬을 직접 예측하도록 한다. 본 연구에서는 두 번째 방법으로 학습을 진행하는 과정에 서비스 요청에 대한 예측이 내부적으로 포함될 수 있다고 판단하여, 두 번째 방법으로 진행한다.

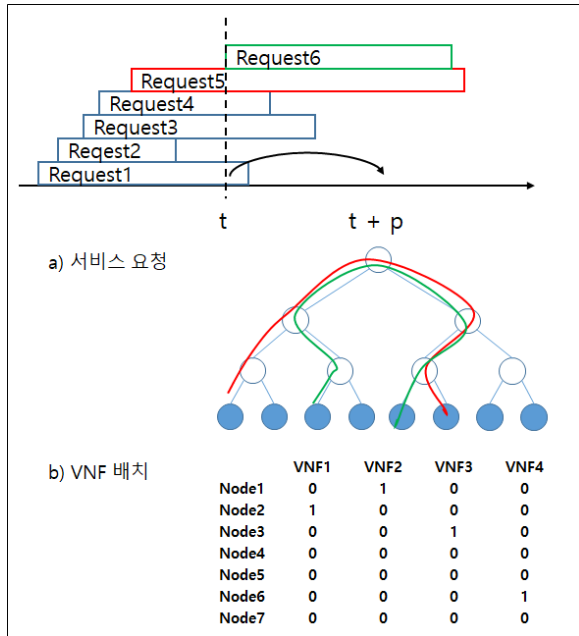


그림 1 VNF 최적 배치 예측 예시

지도 학습 방법으로 VNF 최적 배치 예측 모델을 훈련시키기 위해서, ILP 솔루션을 이용해 시간의 흐름에 따른 서비스 요청 정보를 생성하고, 각 시점에 대한 VNF 최적 배치 정답을 레이블링하는 과정은 다음과 같다. 먼저, MEC (Multi-access Edge Computing) 토폴로지 상에 출발지 노드, 목적지 노드, 도착시간, 지속시간, 트래픽, 최대 지연의 정보를 가지는 서비스 요청 정보를 생성한다. 이를 입력 값으로 하여 CPLEX 프로그램 기반의 ILP 솔루션 [3] 결과로 얻어지는 각 시점에서의 최적의 VNF 배치를 그림 1 의 b)와 같은 매트릭스 형태로 가공한다.

기계학습 모델의 입력 피쳐 (feature)로는 네트워크 토폴로지, 서버의 자원 할당, 네트워크 대역폭 할당,

VNF 타입별 카탈로그와 같은 네트워크 환경 정보뿐만 아니라, 생성한 서비스 요청에 대한 ILP 솔루션의 결과인 VNF 배치 및 SFC 라우팅 정보를 바탕으로 각 노드의 자원 사용량, 대역폭 사용량 등을 계산하여 사용한다. 모델이 예측하는 VNF 배치 결과가 예측 대상이 되는 미래 시점의 ILP 솔루션의 결과와 같아지도록 목적 함수 (objective function)를 구성하여 ILP 솔루션의 VNF 배치 결과와 가까운 결과를 내도록 학습을 진행한다. 시뮬레이션을 통해 얻은 데이터 중 훈련 데이터에 포함되지 않은 테스트 데이터로 모델의 정확도를 평가한다.

다음은 전 단계에서 도출한 모델의 성능을 실제와 가까운 환경에서 검증하기 위해 시뮬레이션에서 가정한 토폴로지와 같은 테스트베드 (testbed)의 네트워크 환경을 구성한 후, 동일한 서비스 요청을 기준으로 트래픽을 생성한다. 시뮬레이션에서는 ILP 를 통해 생성하는 각 노드의 자원 사용량, 대역폭 사용량 정보를 테스트베드의 각 노드 서버에서 실제로 측정하여 입력 피쳐로 수집한다. 각 시점에 대해 VNF 최적 배치로 예측된 결과를 적용하였을 때와 ILP 솔루션의 결과로 도출된 VNF 배치를 적용하였을 때, VNF 처리 시간 (VNF processing time)과 패킷 누락 비율 (packet drop rate) 등의 QoS (Quality of Service)를 비교함으로써 예측 모델의 성능을 평가한다.

최종적으로 도출되는 VNF 배치 결정 모델은 현재 시점까지 이어지는 단위 구간 (time window)의 서비스 요청 정보와 그에 따라 변화하는 각 노드의 자원 사용량, 대역폭 사용량 등의 정보를 모니터링한다. 이를 바탕으로 5 분 이상 미리 VNF 최적 배치를 결정하여, QoS 성능이 ILP 솔루션 대비 80% 수준을 달성하는 것을 목표로 한다.

IV. 결론

본 논문에서는 MEC 토폴로지 상에서 네트워크의 다양한 상태 정보를 바탕으로 미래 시점의 VNF 최적 배치를 예측하는 기계학습 모델을 제안한다. 제안하는 모델은 기존 VNF 타입 별 수량 증감 결정에 대한 연구 [4]를 개선하여 VNF 타입 별 최적 배치를 네트워크 토폴로지 상 노드의 위치까지 구체적으로 예측한다. 최종적으로는 테스트베드에서 검증된 모델을 적용하여 VNF 배치를 최적으로 결정하는 시스템을 구축하는 것을 목표로 한다.

ACKNOWLEDGMENT

이 논문은 2020 년도 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2018-0-00749, 인공지능 기반 가상 네트워크 관리기술 개발).

참고 문헌

[1] B. Li, W. Lu, S. Liu, and Z. Zhu, "Deep-learning-assisted network orchestration for on-demand and cost-effective VNF service chaining in inter-DC elastic optical networks", IEEE/OSA Journal of Optical Communications and Networking, vol. 10, no. 10, pp. D29-D41, Oct. 2018.

[2] X. Zhang, C. Wu, Z Li and F. C. M. Lau, "Proactive VNF Provisioning with Multi-timescale Cloud Resources: Fusing Online Learning and Online Optimization", in 2017 IEEE Conference on Computer Communications, pp. 1-9, Oct. 2017.

- [3] M. F. Bari, S. R. Chowdhury, R. Ahmed, and R. Boutaba, "On Orchestrating Virtual Network Functions", in 11th International Conference on Network and Service Management, pp. 50-56, Nov. 2015.
- [4] S. Lange, et al., "Predicting VNF Deployment Decisions under Dynamically Changing Network Conditions", in 2019 15th International Conference on Network and Service Management, pp.1-9, Oct. 2019.