NFV 관리를 위한 VNF 이상 탐지 시스템에 대한 연구

홍지범, 박수현, 유재형, 홍원기

포항공과대학교 컴퓨터공학과

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

A Study on VNF Anomaly Detection System for NFV Environment Management

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

Department of Computer Science and Engineering, POSTECH

요 약

Software-Defined Networking (SDN) 및 Network Function Virtualization (NFV)의 개념이 제안된 이후, 현재 통신 사업자 및 서비스 제공업체는 SDN/NFV 기술을 활용하여 기존의 서비스를 보다 효율적으로 제공하기 위해 노력하고 있다. 하지만 데이터 센터에서 운용되는 가상 네트워크가 점점 복잡해짐에 따라 리소스 할당, 장애 관리 등과 같은 다양하고 새로운 네트워크 관리 문제가 발생하게 되었다. 복잡해지는 관리 문제를 해결하기 위해서는 우선 가상 네트워크에서 동작하는 Virtualized Network Function (VNF)의 리소스 사용량 및 네트워크 트래픽 로드를 모니터링하고 분석할 필요가 있다. 이를 위해 본 논문에서는 NFV 환경에서 시스템 리소스 및 트래픽 과부하로 인해 발생하는 Service Level Agreement (SLA) 위반과 관련된 VNF 의 이상 상태를 탐지하는 시스템을 제안한다. 제안하는 시스템은 실제 테스트베드에서 수집한 데이터로 머신러닝 (Machine Learning) 모델을 학습시켜 가상 네트워크 환경에서 동작하는 VNF 들의 이상 상태 (anomaly) 탐지에 활용할 수 있다.

Ⅰ. 서 론

오늘날 Software-Defined Networking (SDN) 및 Network Function Virtualization (NFV) 기술은 네트워크 가상화를 통해 네트워크를 구축하고 운영하는 새로운 방법을 제공하고 있다. SDN/NFV 기술은 기존 하드웨어 중심의 폐쇄된 네트워크 기능을 소프트웨어 형태의 Virtualized Network Functions (VNFs)로 대체하여 비용을 절감시킨다. 또한 클라우드 컴퓨팅 (Cloud Computing)은 이러한 기술들을 활용하여 데이터 센터에서 소요되는 컴퓨팅 리소스를 보다 효율적으로 사용하고, 다양한 애플리케이션 서비스를 유연하고 효율적으로 배포 하는 것을 가능하게 한다 [1].

이에 따라 통신 사업자와 서비스 제공업체는 최근 SDN/NFV 및 클라우드 컴퓨팅 기술을 이용하여 다양한 애플리케이션 서비스를 운영하거나 가상 환경에서 Evolved Packet Core (EPC), IP Multimedia Subsystem (IMS)과 같은 핵심 네트워크 기능을 구현하고 있다. 이를 위해 통신 사업자 및 서비스 제공업체는 가상 네트워크에 가상 머신 (Virtual Machine) 또는 컨테이너 (container)를 설치한 후, 원하는 서비스를 동작시킨다.

하지만 기존 하드웨어 기반의 네트워크 장치 운용과는 다르게 NFV 환경에서는 가상 서버 또는 가상 네트워크처럼 자원들을 가상화 하여 서비스 운용이 이루어지기 때문에 가상 네트워크가 점점 복잡해진다는 단점이 존재한다. 이로 인해 리소스 할당 최적화, 장애 관리 등과 같은 다양하고 새로운 네트워크 관리 문제가 발생하게 되어 서비스 및 시스템의 심각한 장애를 유발시킬 수 있다. 따라서 가상 네트워크에서 동작하는 VNF 의 시스템 리소스 사용량 및 네트워크 트래픽 로드와 같은 동작 상태를 분석할 필요가 있다.

본 논문에서는 가상 네트워크의 관리 문제를 해결하기 위한 방법의 일환으로, 머신러닝 (Machine Learning)을 통해 Service Level Agreement (SLA) 위반과 관련된 지표를 학습시켜 서비스를 제공하는 VNF 의 이상 상태를 실시간으로 탐지 (Anomaly Detection)하는 시스템을 제안한다. 그리고 제안하는 이상 탐지 시스템을 활용한 향후 연구 방향을 제시한다.

Ⅱ. 관련 연구

VNF 의 이상 탐지는 CPU, 메모리 사용량과 같은 측정치 (metric)의 임계값 (threshold)을 이용하여 탐지할 수 있지만 이는 단순히 한 측정치의 임계값을 기반으로 이상 여부를 판단하기 때문에 많은 오탐 (false alarm)을 유발한다. 따라서 기존의 이상 탐지 연구는 대체로 시계열 (time-series) 데이터 및 통계 알고리즘을 이용하여 VNF 및 시스템의 상태를 판단하고 있다. STL (Seasonal Trend Decomposition using LOESS) 알고리즘을 적용하여 이상 상태를 탐지하는 연구 [2]는 클라우드 환경에서 수집한 시계열 CPU 데이터의 계절성 요인을 함께 고려하여 3-Sigma 규칙에 의해 정의된 임계값을 넘는 측정치를 이상 상태로 탐지한다.

최근에는 머신러닝과 같은 인공지능 기술을 네트워크 관리에 적용하려는 연구가 활발히 진행되고 있다. 선행연구 [3]는 vIMS 환경에서 동작하는 컨테이너화된 (containerized) VNF 의 이상 상태를 지도학습 (supervised learning) 기반의 Random Forest 알고리즘

을 이용하여 탐지하는 방법을 제안하고 있다. 또한 다른 머신러닝 알고리즘을 적용한 선행 연구 [4]에서는 Support Vector Machine (SVM), Decision Tree (DT), Neural Network (NN) 알고리즘을 통해 VNF 의 성능 이상을 탐지하고 각 학습 모델의 성능을 비교하고 있다.

Ⅲ. 머신러닝 기반의 VNF 이상 탐지 시스템

기존 VNF 이상 탐지 시스템은 대부분 CPU 과부하, 메모리 부족 등의 상태만을 이상으로 판단한다는 한계가 존재한다. 본 논문에서는 리소스 및 네트워크 사용량을 포함하여 SLA 위반과 관련된 지표를 기반으로 서비스의 상태를 반영함으로써 보다 정확하게 이상 상태를 탐지하는 시스템을 제안한다. 제안하는 머신러닝 기반의 VNF 이상 탐지 시스템의 구조는 그림 1 과 같다.

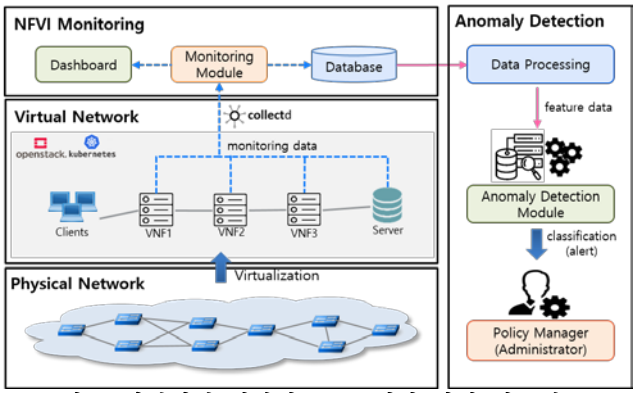

그림 1 머신러닝 기반의 VNF 이상 탐지 시스템 구조

먼저 하드웨어를 통해 구성된 물리 네트워크 환경에서 OpenStack 및 Docker/Kubernetes 와 같은 Virtual Infrastructure Manager (VIM)을 이용하여 NFV Infrastructure (NFVI)를 구축한다. 이러한 NFVI 환경은 인공지능 기반의 NFV 관리 플랫폼 선행 연구 [5]를 기반으로 이루어지며, 다양한 서비스를 제공하기 위해 구성된 가상 네트워크에서 각 서비스에 부합하는 VNF 들을 동작시킨다. 운영중인 가상 네트워크에 대한 데이터는 각 VNF 내부에 설치된 모니터링 에이전트인 Collectd 에서 CPU, memory, disk I/O 와 같은 리소스 사용량, 네트워크 트래픽 로드 등과 같은 네트워크 데이터를 실시간으로 수집하여 모니터링 모듈로 보낸다. 모니터링 모듈은 전달받은 데이터를 대시보드에 전달하여 실시간으로 모니터링 데이터를 시각화하여 보여주거나 데이터베이스 플랫폼에 보내어 시계열 형태의 데이터로 저장한다.

다음으로 머신러닝을 통해 사전 학습하여 도출된 모델 기반의 이상 탐지 모듈은 데이터베이스에서 주기적으로 데이터를 가져온 후, VNF 의 이상 상태 탐지를 위한 분류에 사용하기 위해 데이터를 feature 데이터로 변환시킨다. 이렇게 변환된 feature 데이터를 기반으로 이상 탐지 모듈은 실시간으로 가상 네트워크에서 동작 중인 VNF 들의 현재 상태를 정상 및 이상 상태로 분류하여 판단한다. 이 때 이상 탐지 모듈이 특정 VNF 의 상태에 문제가 있다고 판단을 하면 이상 탐지 모듈은 NFV 환경 관리를 위한 정책을 세우는 Policy Manager 혹은 네트워크 관리자에게 경보 (alert)를 보내어 운영 중인 서비스에 이상 징후가 발생했음을 알린다. 이와 같은 머신러닝 기반의 VNF 이상 탐지 시스템을 구축하기 위해 머신러닝 알고리즘을 통해 이상 탐지 모델을 학습시키는 과정이 필요하다. 가상 네트워크에서 수집한 데이터를 통해 VNF 이상 탐지 모델을 학습시키는 과정은 그림 2 와 같다.

NFVI 환경에서 동작 중인 가상 네트워크를 모니터링하여 데이터를 수집 및 저장하는 방법은 앞선 방법과 동

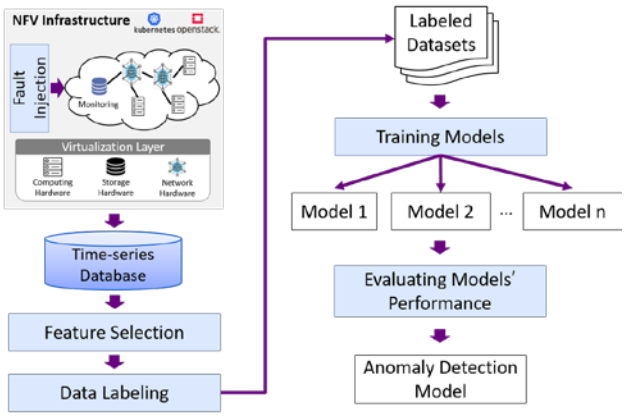

그림 2 VNF 이상 탐지 모델의 학습 과정

일하며, 이를 통해 CPU, memory, disk I/O, 트래픽 로드 등과 같은 다양한 측정치를 세분화하여 수집한다. 하지만 가상 네트워크의 운영 데이터를 수집할 때 실제 네트워크 환경에서는 비정상적인 동작을 나타내는 이상 상태가 거의 발생하지 않으므로 결함 주입 (fault injection)을 통해 CPU 과부하, 메모리 부족, 트래픽 과부하 등의 이상 상태를 발생시켜 SLA 를 위반하도록 유도한다.

이렇게 수집한 모니터링 데이터는 시계열 데이터베이스에 저장된다. 저장된 데이터는 지도학습 기반의 머신러닝 알고리즘 학습에 활용할 수 있도록 수집한 데이터의 여러 측정치 중에서 이상 상태 분류에 핵심이 되는 feature 를 추출하며, 데이터 레이블링은 결함 주입으로 패킷 손실률 (packet loss rate), 지연 (latency), 서비스 요청 실패율 (service request failure rate) 등과 같은 SLA 위반이 발생했을 시점의 데이터를 이상 상태로 레이블링하여 정상 및 이상 상태의 데이터셋을 생성한다.

다음으로 레이블링된 데이터셋을 통해 지도학습 기반의 머신러닝 알고리즘을 각각 학습시킨다. 학습을 위한 환경은 머신러닝/딥러닝 프레임워크인 TensorFlow 및 Pytorch 를 이용한다. 학습에 사용되는 알고리즘은 앞서 관련 연구에서 언급한 SVM, DT, Random Forest 와 같은 알고리즘과 함께 Gradient Boosting Machine (GBM), eXtreme Gradient Boost (XGBoost) 등과 같이 최근 뛰어난 성능을 보여주는 알고리즘을 사용하여 학습시킨다.

마지막으로, 각 알고리즘들을 기반으로 학습된 이상 탐지 모델들의 성능을 accuracy, F1-Score 등을 이용하여 비교한 후, 가장 좋은 성능을 보이는 모델을 제안하는 시스템에서 채용하는 이상 탐지 모듈의 기반 모델로 도출하여 이상 탐지 시스템을 구현한다.

Ⅳ. 결론 및 향후 연구

본 논문에서는 NFV 환경 관리를 위한 머신러닝 기반의 VNF 이상 탐지 시스템을 제안하였다. 제안하는 방법은 가상 네트워크 환경에서 데이터를 수집하여 지도학습 기반의 머신러닝 알고리즘을 통해 이상 탐지 모델을 학습시키고, 이를 SLA 위반 여부를 바탕으로 제공 중인 서비스의 이상 징후를 탐지함으로써 기존 이상 탐지 시스템을 개선하여 서비스의 심각한 장애가 발생하기 전에 네트워크를 선제적으로 관리할 수 있다.

향후 연구로 제안하는 방법을 바탕으로 가상 네트워크의 데이터를 생성 및 수집한 후, 다양한 머신러닝 알고리즘의 성능을 비교하여 최적의 성능을 보이는 VNF 이상 탐지 모델을 도출한다. 또한 도출된 모델을 바탕으로 VNF 이상 탐지 시스템을 구현한다. 마지막으로, 실제 환경에서의 실험을 통해 제안하는 이상 탐지 시스템의 성능을 검증한다.

ACKNOWLEDGMENT

이 논문은 2020 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2015-0-00575, 글로벌 SDN/NFV 공개소프트웨어 핵심 모듈/기능 개발, 2018-0-00749, 인공지능 기반 가상 네트워크 관리기술 개발).

참 고 문 헌

[1] O. Sefraoui, M. Aissaoui, and M. Eleuldj, "Openstack: toward an open-source solution for cloud computing," International Journal of Computer Applications, vol. 55, no. 3, pp. 38-42, Oct. 2012.

[2] J. Hochenbaum, O. S. Vallis, and A. Kejariwal, "Automatic anomaly detection in the cloud via statistical learning," Computing Research Repository, abs/1704.07706, 2017.

[3] C. Sauvanaud, K. Lazri, M. Kaaniche, and K. Kanoun, "Anomaly detection and root cause localization in virtual network functions," In 2016 IEEE 27th International Symposium on Software Reliability Engineering, pp. 196-206, Oct. 2016.

[4] J. Qiu, et al., "Performance anomaly detection models of virtual machines for network function virtualization infrastructure with machine learning," In Artificial Neural Networks and Machine Learning - ICANN 2018, pp. 479-488. Springer International Publishing, Oct. 2018.

[5] 정세연, 이도영, 유재형, 홍원기, "인공지능 기반 NFV 관리 플랫폼," In KNOM Conference 2019, pp. 40-42, May 2019.