

# SDN 기반 네트워크 환경에서 기계학습을 이용한 네트워크 장애분석 연구

강효성  
KT 연구개발센터

kang.hyosung@kt.com

## A Study on Network Failure Analysis using Machine Learning in SDN-based Network Environment

Kang Hyosung  
KT

### 요 약

일반적인 네트워크 환경에서 장애가 발생할 경우 빠른 장애발생 여부 판단과 그 원인 파악이 어려운 문제, 장애복구와 감지까지 많은 시간이 걸린다는 점, 그리고 대부분 사전 예방보다 사후 장애복구에 집중된다는 문제점들이 존재했었다. 그리고 이런 문제점들을 개선할 수 있었던 것이 바로 SDN 이다. SDN 을 통한 중앙집중적인 네트워크 관리 방식을 통해 네트워크 장애 발생시 빠른 장애감지와 복구가 가능해졌지만 SDN 기반 네트워크 환경에서도 일부 장애의 경우 여전히 네트워크 관리자가 직접 장애발생지점으로 가서 현장점검이 필요하다 보니 사전 예방보다 사후 장애복구에만 해결방법이 집중되었다. 그래서 본 논문에서는 네트워크 장애에 대해서 SDN 의 장점을 그대로 가져가면서도 앞의 문제점들을 개선하기 위해 기계학습 방식을 적용하여 네트워크 장애가 발생할 수 있는 잠재적 장애 포인트를 예측하고, 사후 장애복구가 아니라 사전 장애 예방이 되도록 지원하는 해결방법을 제안하고자 한다.

### I. 서 론

기본적으로 네트워크 통신의 장애에는 발생위치를 기준으로 크게 2 가지로 구분된다. 하나는 네트워크의 유무선 전송채널에서 발생한 장애이고, 다른 하나는 네트워크 장비자체에 문제가 발생한 네트워크 장애이다. 일반적인 네트워크 환경에서는 장애가 발생할 경우 장애를 발생한 위치를 파악하는데도 오래 걸릴 뿐만 아니라 다시 장애복구를 위해 네트워크 관리자가 직접 장애가 발생한 네트워크 장비 혹은 전송채널을 교체하거나 점검하여야 한다. 반면 SDN 기반 네트워크 환경에서는 네트워크가 Data Plane 과 Control Plane 이 분리되어 있기 때문에 Control Plane 역할을 수행하는 하나의 SDN 제어기가 하위의 SDN 스위치들을 모니터링하면서 빠르게 장애 감지 및 장애가 발생한 위치 파악이 가능하다[1]. 경우에 따라 이중화를 통해 네트워크 관리자가 직접 전송채널과 장비 교체없이도 원격에서 장애 복구가 가능하다. 하지만 일부 장애의 경우 여전히 네트워크 관리자가 직접 장애발생지점으로 가서 현장점검이 필요하다 보니 사전 예방보다 사후 장애복구에만 해결방법이 집중되었다. 결국 이를 해결하기 위해 네트워크 장애를 진단하고 관리할 수 있는 별도 관리 시스템의 필요성이 더욱 커졌다[2]. 그래서 본 논문에서는 네트워크 장애에 대해서 SDN 의 장점을 그대로 가져가면서도 앞의 문제점들을 개선하기 위해 SVM 분류학습 알고리즘을 적용한 기계학습 방식을 통해 네트워크 장애가 발생할 수 있는 잠재적 장애 포인트를 예측하고, 사후 장애복구가 아니라 사전 장애 예방이 되도록 지원하는 해결방법을 제안하고자 한다.

### II. 제안사항

본 논문에서는 SDN 기반 네트워크 환경에서 SVM 분류알고리즘을 적용하여 분류 예측 방법의 지도학습을

통해 잠재적으로 네트워크 장애 발생가능성이 높은 SDN 스위치를 분류하고 이를 통해 네트워크 장애에 대한 사전 예방을 지원하는 장치들을 제안한다. 해당 장치에는 SVM 분류알고리즘에 적용하기 위한 Data Set 부터 SDN 제어기를 통해 수집된 데이터를 활용한 학습데이터의 생성과 학습이 완료된 후 SDN 스위치의 장애발생여부 검토 및 정확한 학습을 위한 오차율 계산방식까지 모든 관련 장치를 포함한다. 단, 여기서는 SVM 분류모델을 기계학습 알고리즘의 예로 하여 설명하나, 반드시 이와 같이 한정되는 것은 아니며, 성능의 개선이 있다면 다른 분류모델 알고리즘의 적용이 가능하다.



그림 1. 네트워크 장애분석 장치 구조

그림 1 은 본 논문에서 제안하는 장애분석장치의 구성을 나타낸 것이며, 크게 데이터 수집부, 학습데이터 생성부, SVM 학습부, 판단부 이렇게 4 가지 세부장치로 구성된다.

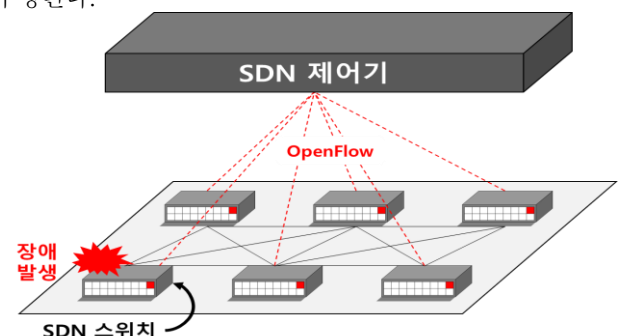


그림 2. SDN 을 통한 스위치의 상태정보 수집과정

데이터 수집부는 2 개의 (일반 및 장애)DB 를 가지고 있으며 그림 2 처럼 각 DB 에는 SDN 제어기를 통해 수집된 SDN 스위치의 상태정보가 장애 발생여부에 따라 나뉘어 그림 3 과 같은 형태의 데이터가 저장되어, 학습데이터 생성을 위한 전처리 역할을 수행한다.

데이터유형	내용	단위	정규화방법
X1	포트사용률	%	$X1 \div 100$
X2	CPU 사용률	%	$X2 \div 100$
X3	Flow 사용률	%	$X3 \div 100$
X4	패킷사이즈	Byte	$X4 \div X4 \text{ 최대값}$
X5	메모리사용률	%	$X5 \div 100$
X6	트래픽처리량	bps	$X6 \div X6 \text{ 최대값}$
X7	내부온도	℃	$X7 \div X7 \text{ 최대값}$

그림 3. 수집부를 통해 수집되는 데이터 정보

그리고 나서 학습데이터 생성부는 데이터 수집부를 통해 저장된 DB 의 데이터 중에서 그림 3 과 같이 장애가 없는 경우는 출력값을 0 을 설정하고 데이터의 30%, 장애가 발생된 경우 출력값을 1 로 설정하고 데이터의 70%를 임의선택해 학습데이터 1000 개를 생성한다.

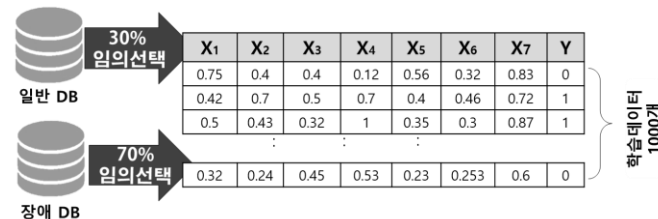


그림 4. DB 데이터를 기반으로 학습데이터의 생성 과정

SVM 학습부는 내부에 SVM 분류모델 알고리즘이 탑재되어, 학습데이터 생성부를 통해 만들어진 그림 4 의 학습데이터를 가지고 SVM 의 학습을 수행한다. 마지막으로 판단부는 그림 5 의 학습오차율 공식을 적용하여 SVM 의 학습 결과에 대한 오차율을 계산하고, 오차율을 줄이기 위한 학습횟수를 결정한다. 여기서는 오차율이 10% 미만인 경우에 학습이 완료되었다고 판단하고 학습을 종료한다. 오차율이 10% 이상인 경우에는 10%미만이 될때까지 재학습을 수행한다.

$$ErrorRate(\%) = \frac{\sum_{k=1}^{100} |f(X_k) - Y|}{100} \quad \text{만약 } f(X_k) = Y \text{ 이면, } ErrorRate = ErrorRate + 1$$

그림 5. SVM 학습오차율 공식

SVM 학습이 완료되면 그림 6 과 같이 8 차원으로 표현된 학습데이터들 사이에서는 네트워크 장애영역과 장애 미발생 영역을 구분하는 판별경계 함수가 결정된다. 해당 판별경계함수는 입력된 SDN 스위치의 상태정보 데이터에 대해서 장애 발생영역에 포함되는 데이터일 경우 1 을 출력하고, 장애 미발생영역에 포함되는 데이터일 경우 0 을 출력하게 된다.

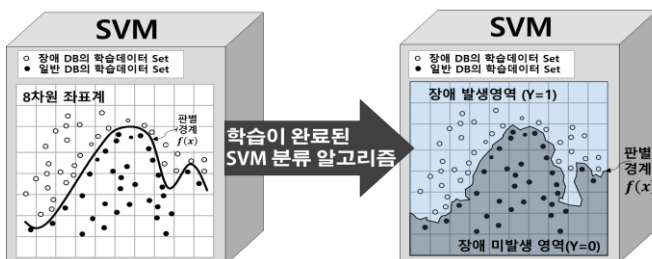


그림 6. SVM 학습결과 판별경계 생성과정

그리고 해당 판별경계함수를 활용하여 그림 7 과 같이 SDN 제어기는 매 일정 주기단위로 SDN 스위치들의 상태정보를 수집하며 해당 판별경계함수에 대입함으로써 나오는 출력값을 가지고 어디에 있는 어떤 스위치 장비가 네트워크 장애 발생 가능성이 높은지 그 가능성을 실시간으로 미리 예측할 수 있다.

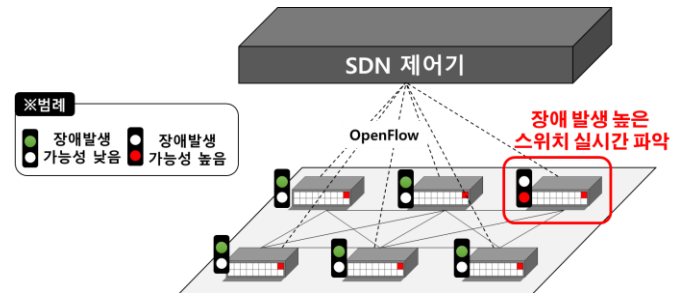


그림 7. 실시간 네트워크 장비의 장애발생 가능성을 검토

### III. 결론

기존 네트워크 모니터링은 장애발생여부를 실시간으로 파악이 어렵고 복구하는데 많은 시간이 소모되었다. 그리고 장애를 사전 예방하는 것이 아니라 사후 조치의 성격이 강했기 때문에 네트워크 통신에서 최상의 QoS 제공에도 한계가 있었다. 하지만 본 연구는 SDN 을 적용함으로써 실시간으로 네트워크 장비들의 상태정보의 모니터링과 위치 관리가 가능해졌고, 네트워크 장애데이터를 기계학습 알고리즘의 학습데이터로 활용하여 네트워크 장비들의 잠재적 장애 발생가능성을 검토해 사후 조치가 아닌 사전 예방 형태의 네트워크 장애에 대한 QoS 보장이 가능하다. 또한 네트워크 관리에 사람의 개입을 줄임으로써 사람의 실수로 인한 네트워크 장애 발생율을 줄이고, 네트워크 관리에 대한 유지비 절감도 가능하다..

### 참 고 문 헌

- [1] Jose Ordonez-Lucena, pablo Ameigeiras, Diego Lopez, Juan J. Ramos-Munoz, javier Lorca, Jesus Folgueira, "Network Slicing for 5G with SDN/NFV: Concepts Architectures and Challenge" IEEE Communications Magazine, Volume.55 , Issue.5 , May.2017.
- [2] D.Johnson, RFC 1297 - NOC Internal Integrated Trouble Ticket System Functional Specification wishlist, Jan 1992. (<http://www.nist.gov/aes>).