

NetConfigQA: 네트워크 설정 해석 능력 평가를 위한 질의응답 데이터셋

Yujin Park

Undergraduate student

Dept. of Information and Communication Engineering
Hanbat National University, Korea

Outline

- 1. Introduction**
- 2. Related work**
- 3. Proposed Method**
- 4. Experiment**
- 5. Conclusion**

1. Introduction

2. Related work

3. Proposed Method

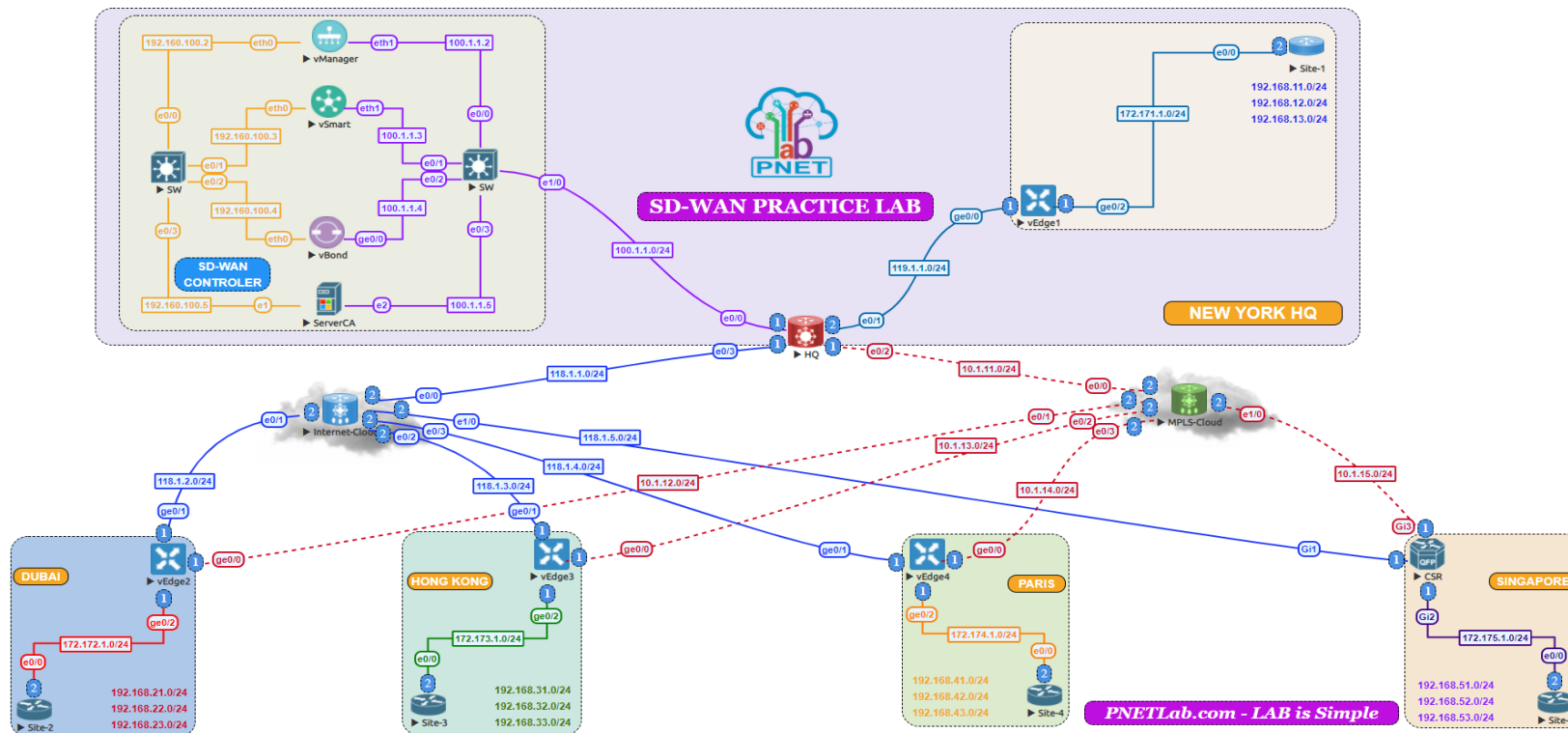
4. Experiment

5. Conclusion

1. Introduction

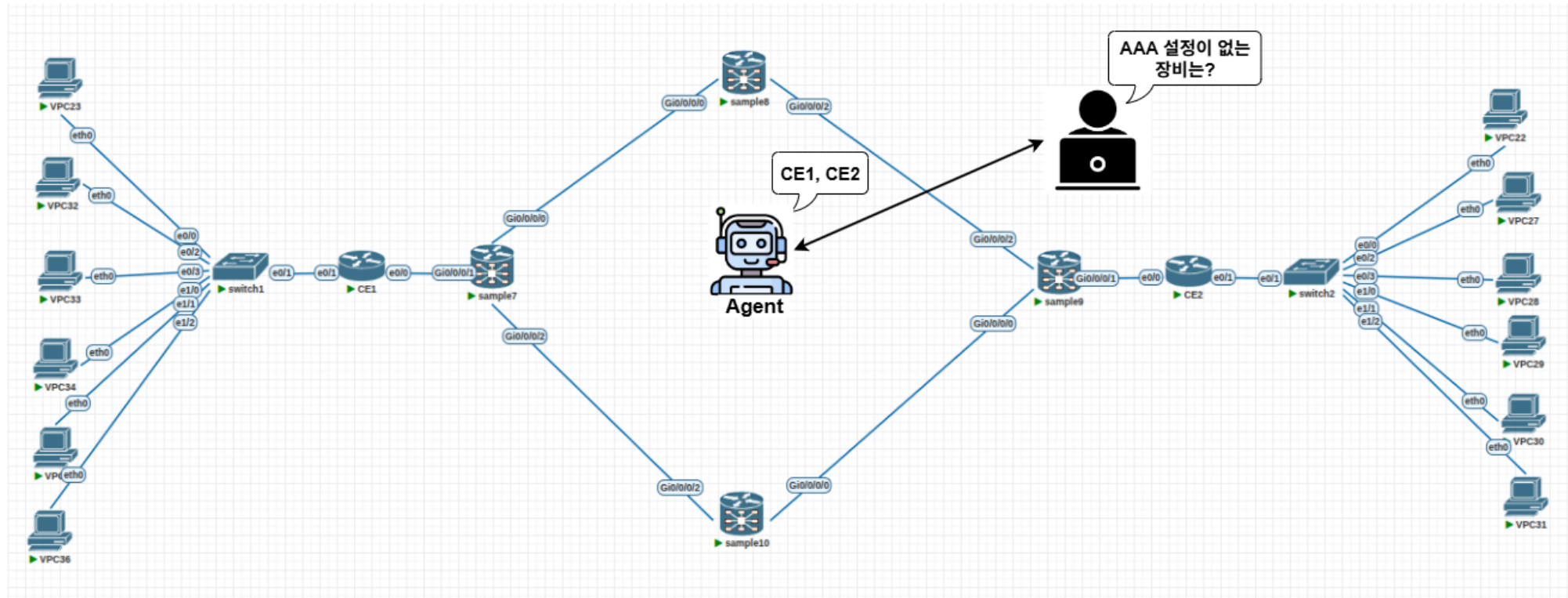
Introduction

- 대규모·멀티벤더 환경으로 네트워크 복잡성 급증
- 스크립트 기반 자동화는 동적 변화에 취약, 장애 예측 및 원인 분석에 한계
- LLM이 대안이 될 수 있음



Introduction

- 대규모, 멀티벤더 환경으로 네트워크 복잡성 급증
- 스크립트 기반 자동화는 동적 변화에 취약, 장애 예측 및 원인 분석에 한계
- LLM이 대안이 될 수 있음



1. Introduction

2. Related Work

3. Proposed Method

4. Experiment

5. Conclusion

2. Related Work

Related Work

- 해석 중심의 네트워크 관리
 - 대부분의 연구가 '생성'에 집중하지만, 이미 구축된 설정을 '해석'하고 '파악'하는 것이 더 중요

벤치마크	주요 목표	입출력
NetConfEval [1]	설정 · 코드 생성	요구→설정/코드/API
NETPRESS [5]	에이전트 운영	지시→액션
NETLLMBENCH [6]	명령 생성	지시→JSON(단일)
NeMoEval [2]	토폴로지 분석코드	질의→그래프코드
NetConfigQA (Ours)	설정 해석 · QA	네트워크 설정→질문, 답

Related Work

- 해석 중심의 네트워크 관리
 - LLM의 네트워크 관리 도메인 이해 능력을 평가하기 위한 질의응답 데이터셋이 필요

sample9 장비와 eBGP로 연결된 피어들의 IP 주소 목록을 알려주세요.	192.168.3.11
AS 65000의 iBGP Full-Mesh 구성은 완벽합니까? (True/False)	TRUE
AS 65001의 iBGP Full-Mesh 구성은 완벽합니까? (True/False)	TRUE
AS 65003의 iBGP Full-Mesh 구성은 완벽합니까? (True/False)	TRUE
AS 65000의 iBGP Full-Mesh에서 누락된 링크는 총 몇 개입니까?	0
AS 65001의 iBGP Full-Mesh에서 누락된 링크는 총 몇 개입니까?	0
AS 65003의 iBGP Full-Mesh에서 누락된 링크는 총 몇 개입니까?	0
sample10 장비와 iBGP로 연결된 피어들의 IP 주소 목록을 알려주세요.	1.1.1.1, 2.2.2.2, 3.3.3.3
sample7 장비와 iBGP로 연결된 피어들의 IP 주소 목록을 알려주세요.	2.2.2.2, 3.3.3.3, 4.4.4.4
sample8 장비와 iBGP로 연결된 피어들의 IP 주소 목록을 알려주세요.	1.1.1.1, 3.3.3.3, 4.4.4.4
sample9 장비와 iBGP로 연결된 피어들의 IP 주소 목록을 알려주세요.	1.1.1.1, 2.2.2.2, 4.4.4.4

1. Introduction

2. Related work

3. Proposed Method

4. Experiment

5. Conclusion

3. Proposed Method

Proposed Method

- XML 설정 파일로부터 질문과 정답을 자동으로 생성
- 질문 템플릿으로 질문을 정의
- 특정 네트워크 토폴로지에 특화된 데이터셋 생성 가능

표 2. 플랫폼별 JSON 표준화 예시

IOS

XML: <Ethernet><name>0/0.100</name></Ethernet>

JSON: interfaces.name="Ethernet0/0.100"

IOS-XR

XML: <GigabitEthernet><id>0/0/0/1.100</id></GigabitEthernet>

JSON: interfaces.name="GigabitEthernet0/0/0/1.100"

표 3. 주요 카테고리별 추출 정보

카테고리	추출 정보 수	추출 정보 예시
Security Policy	5	SSH 활성화 장비
BGP Consistency	4	AS iBGP 이웃 장비
VRF Consistency	4	VRF RD 값
L2VPN Consistency	4	L2VPN 페어
OSPF Consistency	3	OSPF 인터페이스
System Inventory	5	장비 호스트네임
Security Inventory	4	SSH 활성화 여부
Interface Inventory	5	장비 인터페이스 개수
Routing Inventory	4	로컬 BGP AS 번호
Services Inventory	7	VRF 이름
Basic Info	14	시스템 호스트네임
Command Generation	17	BGP 요약 확인 명령어
총합	75	

Proposed Method

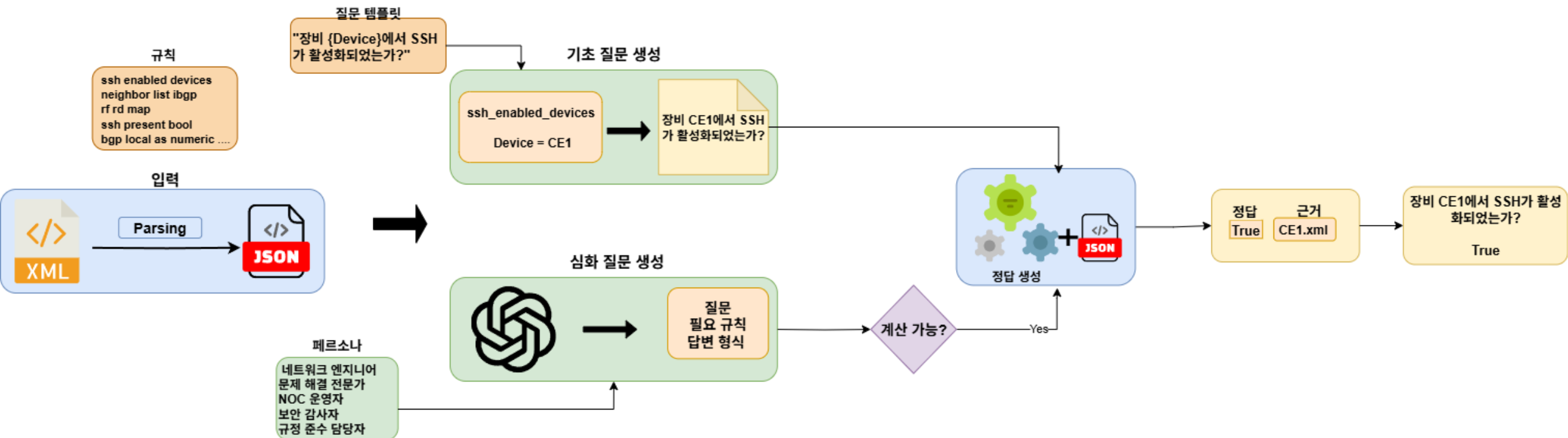
- XML 설정 파일로부터 질문과 정답을 자동으로 생성
- 질문 템플릿으로 질문을 정의
- 특정 네트워크 토폴로지에 특화된 데이터셋 생성 가능

```
# System Inventory
"system_hostname_text": "{host} 장비의 호스트네임은 무엇입니까?",
"system_version_text": "{host} 장비의 운영체제(OS) 버전은 무엇입니까?",
"system_timezone_text": "{host} 장비의 시간대(Timezone)는 무엇입니까?",
"system_user_list": "{host} 장비에 등록된 로컬 사용자 목록을 알려주세요.",
"system_user_count": "{host} 장비에 등록된 로컬 사용자는 총 몇 명입니까?",
# Interface Inventory
"interface_count": "{host} 장비에 설정된 네트워크 인터페이스는 총 몇 개입니까?",
"interface_ip_map": "{host} 장비의 각 인터페이스에 할당된 IP 주소를 알려주세요.",
"interface_vlan_set": "{host} 장비에 설정된 VLAN 목록을 알려주세요.",
"subinterface_count": "{host} 장비에 설정된 서브인터페이스는 총 몇 개입니까?",
"vrf_bind_map": "{host} 장비의 각 인터페이스별 VRF 바인딩 현황을 알려주세요.",
# Device Inventory
```

```
elif metric == "interface_count":
    host = scope.get("host")
    for d in self.devices:
        if host and self._hostname(d) != host: continue
        arr = d.get("interfaces") or []
elif metric == "interface_ip_map":
    host = scope.get("host")
    for d in self.devices:
        if host and self._hostname(d) != host: continue
        mp={}
        for i in d.get("interfaces") or []:
            name=i.get("name") or i.get("id") or ""
            ip =(i.get("ipv4") or i.get("ip") or "")
            if name: mp[name]=ip
        return "map", mp
return "map", {}
```

Proposed Method

- XML 설정 파일로부터 질문과 정답을 자동으로 생성
- 질문 템플릿으로 질문을 정의
- 특정 네트워크 토폴로지에 특화된 데이터셋 생성 가능



Proposed Method

- XML 설정 파일로부터 질문과 정답을 자동으로 생성
- 질문 템플릿으로 질문을 정의
- 특정 네트워크 토폴로지에 특화된 데이터셋 생성 가능

표 6. NetConfigQA 데이터셋 구성

구분	질문 수	비율	평균 토큰	총 토큰
기초 질문	763	93.8%	30.8	23521
심화 질문	50	6.2%	51.8	2589
총계	813	100%	32.1	26110

- 1. Introduction
- 2. Related work
- 3. Proposed Method
- 4. Experiment
- 5. Conclusion

4. Experiment

Experiment

- GPT-4O-mini 모델을 사용한 실험 결과

표 7. 방법론별 성능 비교 결과

평가 지표	Baseline	CoT	Ours
EM	0.666	0.688	0.819
F1	0.707	0.733	0.837
BERTScore	0.878	0.879	0.819

Experiment

- 설정 '해석'을 위한 자동화 QA 파이프라인 **NetConfigQA**
 - 단일 장비/문자열 추출이 아니라 네트워크 전역 속성을 질문함.
 - 정답은 설정/토폴로지로부터 계산된다
 - 모델이 특정 네트워크 망을 얼마나 잘 파악하는지 측정 가능
 - 네트워크 망이 달라져도 변경된 망에 특화된 데이터셋 생성가능

5. Next Plan

Next Plan

- 문제 난이도 상승
 - 단순 사실 확인 → **네트워크 전역을 아우르는 복합 추론**으로 확장
 - 네트워크를 그래프(Graph)로 모델링하고, **가상 시나리오**를 시뮬레이션
 - 다수의 장비 설정을 보고, 생각해야 하는 문제들을 추가

Next Plan

• 예시1

- 가상으로 링크 장애를 시뮬레이션하여 네트워크의 생존성을 검증하는 질문을 추가
 - "PE1-P1 코어 링크가 다운되어도, CE1에서 CE2로 가는 대체 경로가 존재하는가?"
- **구현 방안:** 정답 계산 로직에 그래프 모델을 구축하고, 특정 링크(Edge)를 제거한 후 경로가 존재하는지 재계산하는 로직을 추가.

Next Plan

• 예시2

- 네트워크 전체에 걸친 보안 및 라우팅 정책이 올바르게 동작하는지 검증하는 질문을 추가
 - "외부에서 내부 DB 서버로 들어오는 모든 트래픽은 **반드시** 방화벽(FW-1)을 통과하는가?"
- 여러 개의 경유지(Waypoint)를, **정해진 순서대로** 통과하는지를 검증

Next Plan

- 예시3

- 비대칭 라우팅을 검증하는 질문을 추가
 - "서버 A(10.1.1.1)와 클라이언트 B(20.2.2.2) 사이의 통신에서, A에서 B로 가는 트래픽의 경로와 B에서 A로 돌아오는 트래픽의 경로가 서로 다릅니까? 만약 다르다면, 왕복 경로에서 서로 다른 라우터는 무엇입니까?"
- 트래픽이 갈 때와 올 때 다른 경로를 사용하는 **비대칭 라우팅**을 탐지

Next Plan

- 현재 단계에서는 심화 질문을 LLM 에이전트로 생성했지만 재현성과 유지보수를 위해 좋은 질문들을 스크립트 기반 생성으로 전환
- 실제 현업에서 발생하는 질문, 검증 해야하는 항목들을 질문으로 생성할 것

Q & A

Yujin Park

Undergraduate student

Dept. of Information and Communication Engineering
Hanbat National University, Korea

Conclusion

Appendix

- 1. Introduction
- 2. Related work
- 3. Proposed Method
- 4. Experiment
- 5. Conclusion

5. Conclusion