

NS-3 시뮬레이션 분석을 위한 LLM 기반 자연어 쿼리 시스템

김용수*, 남대현, 조효준, 문수빈, 강민교, 이승준, 전형준, 이재우

국립부경대학교 정보통신공학전공

(oloae*, namdh01, dhkdlfemqh, subin23, alsry4706, ubss3231, jun04292569)@pukyong.ac.kr,
jlee0315@pknu.ac.kr

An LLM-based Natural Language Query System for NS-3 Simulation Analysis

Eungsu Kim*, Daehyeon Nam, Hyojun Jo, Subin Moon, Minkyo Kang, Seungjun Lee, Hyeongjun Jeon,
Jaewook Lee

Department Information and Communications Engineering, Pukyong National University

요약

본 논문은 NS-3 네트워크 시뮬레이션 데이터 분석을 위한 자연어 기반 쿼리 인터페이스를 제안한다. 복잡한 쿼리 작성 없이 자연어 질의만으로 시뮬레이션 데이터를 조회할 수 있도록 구현하였다. 제안 시스템은 NS-3 시뮬레이터, Watchdog 기반 실시간 데이터 수집, InfluxDB 시계열 데이터베이스, 그리고 Llama 3.1 기반 Tool-based Agent로 구성된다. 2단계 LLM 호출 구조를 통해 환각을 방지하고 네트워크 도메인 특화 답변을 생성하며, 위치, SINR, 지연시간, 연결 상태, 데이터 전송량 등 6개 카테고리 14개 Tool을 구현하여 다양한 네트워크 지표 조회가 가능하다. 다양한 시나리오 기반의 실험을 통해 제안 시스템의 자연어 질의 처리 성능과 데이터 분석의 정확성을 검증하였다.

I. 서론

NS-3는 LTE/5G 등 다양한 통신 네트워크 환경을 모델링하는 데 널리 사용되는 시뮬레이터이다.[1] 시뮬레이션 결과는 일반적으로 대용량 시계열 데이터베이스인 InfluxDB에 저장되어 사후 분석에 활용된다.[2] 그러나 이러한 데이터를 분석하기 위해서는 복잡한 InfluxQL 쿼리문을 직접 작성하거나, 별도의 Python 데이터 처리 스크립트를 구현해야 하는 어려움이 있다. 이는 네트워크 도메인 지식은 갖추었으나 데이터베이스 쿼리 언어에 익숙하지 않은 연구자들에게 높은 진입 장벽으로 작용하며, 실시간 모니터링 및 장애 원인 파악을 지연시키는 요인이 된다. 최근 대규모 언어 모델(LLM)의 추론 능력(Reasoning)과 외부 도구 사용(Tool Use) 능력을 결합한 연구들이 주목받고 있다. 특히 Yao et al.[3]이 제안한 ReAct(Reasoning and Acting) 프레임워크와 Schick et al.[4]의 Toolformer 연구는 LLM이 단순한 텍스트 생성을 넘어, 외부 API를 호출하여 정확한 정보를 검색하고 판단할 수 있음을 입증하였다. 하지만 이러한 일반적인 LLM 애이전트를 네트워크 시뮬레이션 도메인에 그대로 적용할 경우, 수치 데이터에 대한 부정확한 해석이나 존재하지 않는 데이터를 생성하는 환각(Hallucination) 문제가 발생할 수 있다.

본 논문에서는 이러한 문제를 해결하기 위해 2단계 LLM 호출 구조를 갖춘 자연어 쿼리 인터페이스를 제안한다. 1단계 LLM이 사용자 질문을 분석하여 적절한 데이터 조회 도구를 선택하고, 2단계에서 조회된 실제 데이터를 바탕으로 도메인 지식에 기반한 해석을 수행한다. 이를 통해 사용자는 자연어 질문만으로 시뮬레이션 결과를 즉각적으로 분석할 수 있다.

II. 시스템 구성

제안하는 시스템은 NS-3 시뮬레이터, 데이터 수집 모듈, InfluxDB, 그리고 LLM 애이전트의 4계층으로 구성된다. 먼저 NS-3 시뮬레이터는 5G NR(New Radio) 환경을 모사하는 4gNB mmWave 시나리오를 기반으로 동작하며, 시뮬레이션 과정에서 발생하는 CU(Centralized Unit), DU(Distributed Unit)의 성능 지표와 UE(User Equipment) 위치 정보는 텍스트 로그 파일로 실시간 기록된다.

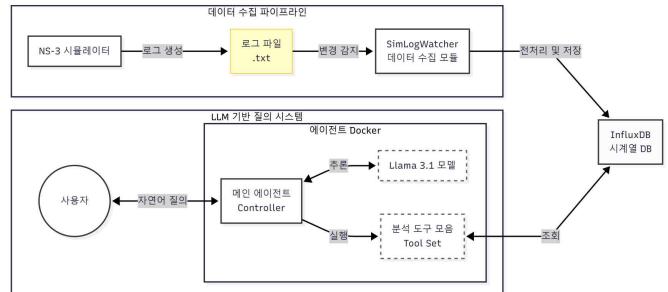


그림 1. 시스템 구성도

이 데이터의 실시간 저장을 위해 본 연구에서는 시뮬레이션 데이터의 실시간 동기화를 위해 Watchdog 라이브러리를 활용한 데이터 수집 모듈을 구현하였다. 이 모듈은 로그 파일의 변경을 감지하고 전처리 과정을 수행하여 InfluxDB에 데이터를 저장한다. 구체적으로 시뮬레이터의 원본 필드명을 적관적인 InfluxDB measurement 이름으로 변환하여 쿼리 가독성을 높이고, PDCP 지연시간과 같은 수치 데이터를 분석에 용이한 단위로 변환하며, 타임스탬프를 데이터베이스 정밀도에 맞춰 동기화한다. 또한 단말 ID(ue_id)와 셀 ID(cell id)를 인덱싱이 적용되는 '태그(Tag)'로 설정하여, 대용량 시계열 데이터 내에서도 애이전트가 특정 조건의 데이터를 즉시 검색할 수 있도록 스키마(데이터 구조)를 최적화하였다.

제안 시스템의 핵심은 환각 방지와 도메인 특화 답변 생성을 위한 2단계 처리 구조이다. 1단계인 도구 선택 단계에서는 사용자가 질문을 입력하면 애이전트가 시스템 프롬프트와 함께 질문을 LLM(Llama 3.1 8B)에 전달한다. 이때 프롬프트는 자연어 답변을 배제하고 오직 JSON 형식으로 도구 선택만 수행하도록 엄격히 제한되어 있어, LLM이 불필요한 텍스트를 생성하여 과정 오류를 일으키는 문제를 방지한다. 이어지는 2단계 답변 생성 단계에서는 애이전트가 앞서 선택된 도구를 실행하여 InfluxDB에서 실제 데이터를 조회하고, 그 결과와 사용자 질문을 분석용 프롬프트와 함께 다시 LLM에 전달하여 최종적인 자연어 답변을 생성한다.

III. 도구 정의 및 프롬프트 설계

카테고리	Tool 이름	기능	주요 파라미터
위치	latest_position	특정 UE 최신 좌표 조회	ue
	position_history	UE 위치 히스토리 조회	ue, limit
	near_xy	특정 좌표 반경 내 UE 검색	x, y, radius
SINR	latest_sinr	특정 UE 최신 SINR 조회	ue, cell
	low_sinr_ues	낮은 SINR(임계값 이하) UE 검색	threshold
지연	latest_latency	PDCP 지연시간 조회	ue, cell
	high_latency_ues	고지연(임계값 이상) UE 검색	threshold
연결	active_ues_per_cell	셀당 활성 UE 수 집계	-
	rrc_connection_time	RRC 연결 시간 조회	cell
전송	throughput_stats	데이터 전송 통계 조회	ue, cell
범용	query_measurement	일반 측정값 쿼리를 조회하는 도구	measurement, limit

표 1. 분석 도구 목록

에이전트가 활용할 수 있는 분석 도구는 표 1과 같이 위치, SINR, 지연 시간, 연결 상태, 전송량, 범용 측정 등 6개 카테고리에 걸쳐 총 14개로 구현되었다. 각 도구는 InfluxQL 쿼리를 캡슐화한 Python 함수 형태로 정의되어, 전처리된 데이터 스키마에 맞춰 소문자 Measurement 이름을 사용한다. 예를 들어 특정 단말의 최신 좌표를 조회하는 latest_position, SINR 이력을 조회하는 sinr_history, 특정 임계값 이상의 지연 시간을 겪는 단말을 검색하는 high_latency_ues 등의 도구가 포함된다.

2단계 LLM 호출 시 사용되는 분석용 시스템 프롬프트에는 네트워크 도메인 전문가의 지식이 사전에 주입되어 있다. 예를 들어 "SINR 값이 10dB 미만일 경우 '불량'으로 판단하고 핸드오버 필요성을 언급하라"거나, "PDCP 지연시간이 50ms 미만이면 '우수'로 평가하라"는 구체적인 해석 가이드라인이 포함된다. 이를 통해 LLM은 단순한 수치 나열에 그치지 않고, "현재 UE 9의 SINR은 -8dB로 통신 상태가 불량하여 인접 셀로의 핸드오버가 권장됩니다"와 같이 사용자가 활용 가능한 분석 결과를 제공할 수 있다.

IV. 실험 결과

구현된 시스템의 실제 동작을 검증하기 위해 "Cell 2에 접속한 UE 번호와 SINR 값을 알려줘"라는 자연어 질의를 입력하였다. 실험 결과, 시스템은 1단계에서 사용자의 의도를 파악하여 latest_sinr 도구를 선택하고 {'cell': '2'} 인자를 정확히 도출하였다. 이후 InfluxDB 쿼리 실행을 통해 획득한 원본 데이터를 바탕으로, 2단계 생성 과정에서 "UE 1: 39.0 dB, UE 6: 97.0 dB" 등 접속된 단말들의 구체적인 신호 품질 수치를 나열하고, "24.0 dB부터 97.0 dB까지 다양하다"는 요약 정보를 포함하여 사용자에게 자연어 형태로 반환하였다.

또한 시스템의 강건함을 확인하기 위해 위치 추적, 신호 품질, 지연 시간, 연결 상태 등 6개 카테고리에 걸쳐 다양한 형태의 질의 테스트를 수행하였다. 실험 결과, 위치 조회, SINR 측정, Latency 확인 등 사용자 의도가 명확한 단일 작업(Single-turn Task)에 대해서는 90% 이상의 높은 도구 선택 정확도를 보였다. 특히 '인터넷이 느리다'와 같은 추상적 표현을 high_latency 도구로 매핑하는 등 우수한 문맥 이해 능력을 확인하였다.

나아가, 단순한 분석의 편의성을 넘어 데이터의 신뢰성 확보 측면에서 본 시스템은 폐쇄형 도메인에 최적화된 성능을 입증하였다. 네트워크

관제와 같이 높은 데이터 정합성이 요구되는 폐쇄형 도메인에서는 일반적인 LLM의 추론만으로는 한계가 있다. 제안 시스템은 외부 지식 없이도 내부 DB의 실측값을 직접 참조하도록 설계되어, 'UE 3의 현재 위치'와 같은 동적 데이터 질의에서도 환각 없이 사실에 입각한 답변을 제공함을 확인하였다.

반면, 'UE 9의 모든 정보'와 같이 복합적인 정보를 요구하는 질의에 대해서는 에이전트가 적절한 도구를 특정하지 못해 오류가 발생하였다. 이는 향후 단계 추론(Multi-step Reasoning) 기능 도입과 데이터 수집 시나리오 확장을 통해 개선할 예정이다.

V. 결론

본 논문에서는 NS-3 네트워크 시뮬레이션 데이터 분석의 진입 장벽을 낮추기 위해, LLM 기반의 자연어 쿼리 인터페이스를 제안하고 구현하였다. 제안 시스템은 2단계 구조의 Tool-based Agent를 도입하여 기존 LLM의 환각 문제를 구조적으로 방지하고, 네트워크 도메인 지식이 주입된 프롬프트 엔지니어링을 통해 전문가 수준의 답변을 생성하도록 설계되었다. 또한 전체 시스템을 Docker 컨테이너 환경으로 구축하여 연구의 재현성과 배포 용이성을 확보하였으며, 비교 실험을 통해 자연어 처리의 정확도와 실무 적용 가능성을 입증하였다. 향후 연구로는 DSPy(Declarative Self-improving Python) 프레임워크를 도입하여 프롬프트 최적화를 자동화하고, 현재의 단일 작업 처리 한계를 극복하기 위해 복합 쿼리 처리가 가능한 단계 추론(Multi-step Reasoning) 및 도구 병렬 실행 메커니즘을 고도화할 계획이다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입니다. (No. RS-2025-00558169)

참고 문헌

- [1] NS-3 Consortium, "ns-3 Network Simulator," <https://www.nsnam.org/>.
- [2] P. Dixit et al., "Enabling Real-Time Analytics in Network Simulators using InfluxDB and Grafana,"
- [3] S. Yao et al., "ReAct: Synergizing Reasoning and Acting in Language Models," International Conference on Learning Representations (ICLR), 2023.
- [4] T. Schick et al., "Toolformer: Language Models Can Teach Themselves to Use Tools," Advances in Neural Information Processing Systems (NeurIPS), 2023.
- [5] H. Touvron et al., "Llama 2: Open Foundation and Fine-Tuned Chat Models," arXiv preprint arXiv:2307.09288, 2023.