

클라우드 애플리케이션의 성능 모니터링을 위한 시계열 데이터 분석 연구

홍두표¹, 김동완¹, 신용태²

¹승실대학교 컴퓨터학과

²승실대학교 컴퓨터학부

envydp99@soongsil.ac.kr, kk5339@soongsil.ac.kr, shin@ssu.ac.kr

A study on time series data analysis for performance monitoring of cloud applications

Dupyoo Hong¹, Dongwan Kim¹, Yongtae Shin²

¹Dept. of Computer Science and Engineering, Soongsil University

²School of Computer Science and Engineering, Soongsil University

요 약

클라우드 애플리케이션의 성능 모니터링 방법에는 클라우드 소프트웨어 스택의 인프라, 플랫폼 및 애플리케이션 계층에서 수집한 시계열 데이터 분석이라는 방법이 존재한다. 클라우드 컴퓨팅 환경에서 운영되는 서비스 간의 런타임 종속성을 분석하는 것은 클라우드 리소스 관리를 수행하기 위해 필요한 단계이다. 본 논문에서는 Bi-LSTM 기법을 활용해 클라우드 애플리케이션의 관계를 분석하고 종속성을 찾아 모니터링 성능을 향상시키는 시스템을 제안한다. 제안하는 시스템은 클라우드 스택의 모든 계층으로부터 시계열 데이터를 수집하여 인공지능 모델을 훈련, 재훈련 및 업데이트 과정을 진행한다. 본 논문에서는 Bi-LSTM 모델을 활용하여 훈련 중에 학습된 성능 메트릭 간의 종속성을 발견한다.

1. 서 론

최근 클라우드 애플리케이션의 성능 모니터링 방법에는 클라우드 소프트웨어 스택의 인프라, 플랫폼 및 애플리케이션 계층에서 수집한 시계열 데이터 분석하는 방법이 존재한다. 클라우드 컴퓨팅 환경에서 운영되는 서비스 간의 런타임 종속성을 분석하는 것은 클라우드 리소스 관리를 수행하기 위해 필요한 단계이다.

그러나 애플리케이션 간의 종속성을 파악하는 것은 애플리케이션 간의 관계를 파악해야 하기 때문에 종속성을 파악하기 어렵다. 이에 본 논문에서는 인공지능 기법을 활용해 애플리케이션 간의 종속성을 분석하여 클라우드 환경에서 모니터링 성능을 향상시킨다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 클라우드 애플리케이션의 종속성 분석에 관한 연구를 살펴본다. 3장에서는 본 논문에서 제안하는 인공지능을 활용한 클라우드 환경에서 애플리케이션 종속성 분석을 통한 모니터링 기법을 제시하고 마지막 4장에서는 결론 및 향후 연구 과제를 제시한다.

2. 관련 연구

본 장에서는 기존 클라우드 애플리케이션의 종속성을 파악하는 인공지능 기법에 대해서 살펴보고 그 결과를 기반으로 제안하는 종속성 분석 기법의 요구 사항을 도출한다.

기존 클라우드 애플리케이션 종속성 분석은 시계열 데이터 분석 알고리즘인 ARIMA와 HoltWinters 모델을 활용한 연구 등을 살펴본다.

2.1 Bi-LSTM

Bi-LSTM(Bidirectional recurrent neural networks)은 양방향 LSTM 모델로 기존 RNN 모델에서 발생하는 기울기 소실 문제를 해결하기 위해 고안된 모델로 특히 시계열 데이터와 같이 시간적 순서가 있는 데이터 처리에 효과적인 모델이다.

Bi-LSTM[2]은 입력 시퀀스를 앞뒤 양방향으로 처리하여 현재 시점에서의 입력에 대한 정보뿐만 아니라 이전과 이후의 입력 정보를 모두 활용하여 출력 값을 생성한다.

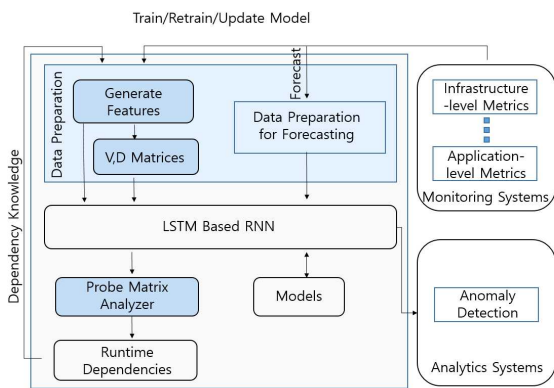
2.2 클라우드 애플리케이션의 종속성 분석 연구

기존의 클라우드 애플리케이션 종속성에 대한 분석

연구는 시계열 데이터 분석 알고리즘인 ARIMA와 Holt-Winters 계절 모델을 활용하여 진행되고 있다. ARIMA 모델은 시계열 데이터 분석을 위한 대표적인 모델 중 하나이다. ARIMA 모델은 자기회귀 (Autoregressive), 누적차분(Integrated), 이동평균 (Moving Average) 모델의 조합으로 이루어져 있다. ARIMA 모델의 주요 개념 중인 자기회귀는 과거의 값이 현재 값에 영향을 미치는 상관관계를 의미이다. Holt-Winters 알고리즘[3]은 시계열 데이터를 예측하는 방법으로 추세(Trend), 계절성(Seasonality), 불규칙성(Irregularity)을 모두 고려하여 예측하는 방법이다. Holt-Winters 알고리즘은 Exponential Smoothing 기법을 사용하여 시계열 데이터를 예측한다. 이를 위해서는 일반적으로 3개의 파라미터를 설정해야 한다.

3. 제안하는 시스템

본 논문에서는 클라우드 네이티브 애플리케이션 간의 종속성을 분석하는 연구를 진행한다. 종속성 분석을 위해 시계열 데이터를 얻을 수 있는 샘플 클라우드 애플리케이션을 구축한다. 측정 데이터셋을 분석하여 수행하도록 클라우드 네이티브 환경을 구성한다.



[그림 1] 제안하는 Bi-LSTM 기반 클라우드 관리 시스템

[그림 1]은 종속성을 추출한 다음 해당 정보를 시계열 예측에 사용하기 위해 LSTM 네트워크를 활용하는 시스템 아키텍처이다. 인프라, 플랫폼 및 애플리케이션을 포함하는 클라우드 스택의 다양한 계층에서 모니터링 시스템의 성능 측정을 위해 수집한다. 모든 계층의 새 모니터링 데이터를 사용하여 모델을

재훈련한다. 네트워크가 훈련하고 훈련 오류가 원하는 최소 수준으로 수렴하면 제안하는 시스템은 Probe Matrix Analyzer에 대한 프로브 벡터 가중치를 사용한다.

4. 결론

본 논문에서는 Bi-LSTM 기법을 활용해 클라우드 애플리케이션의 관계를 분석하고 종속성을 찾아 모니터링 성능을 향상시키는 시스템을 제안한다. 제안하는 시스템은 클라우드 스택의 모든 계층으로부터 시계열 데이터를 수집하여 인공지능 모델을 훈련, 재훈련 및 업데이트 과정을 진행한다. 향후 제안하는 시스템의 현실적인 검증을 위해 설계를 기반으로 구축이 필요하다.

사사문구

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음” (IITP-2023-2020-0-01602)

참고문헌

[1] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735 - 1780, Nov. 1997.

[2] Seung-Hyeon Yoon, Jeong-Min Lee, Ji-Min Oh, Ki Young Lee.(2022).A Judicial Precedent Analysis Service Based on the Natural Language Processing and the Bi-LSTM Algorithm.대한전자공학회 학술대회,(),2217-2218.

[2] K. Kalpakis, D. Gada, and V. Puttagunta, “Distance measures for effective clustering of arima time-series,” in *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on. IEEE, 2001*, pp. 273 - 280

[3] A. M. De Livera, R. J. Hyndman, and R. D. Snyder, “Forecasting time series with complex seasonal patterns using exponential smoothing,” *Journal of the American Statistical Association*, vol. 106, no. 496, pp. 1513 - 1527, 2011.