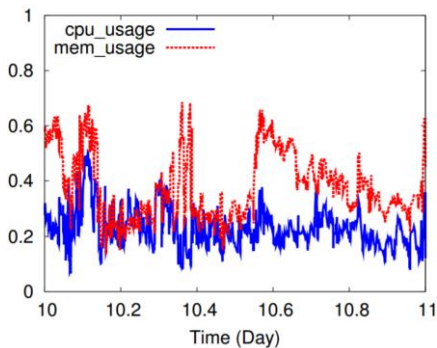


학부생 연구기회 프로그램 (UROP) 공고

◆ 담당교수 : Bernhard Egger	◆ 연구실명 : 컴퓨터 시스템 및 플랫폼
◆ UROP 연구 과제명 : 데이터센터 워크로드의 시계열 예측기법 연구	
◆ 모집대상 : 제한없음	
◆ 모집기간 : ~ 2016년 12월 말	

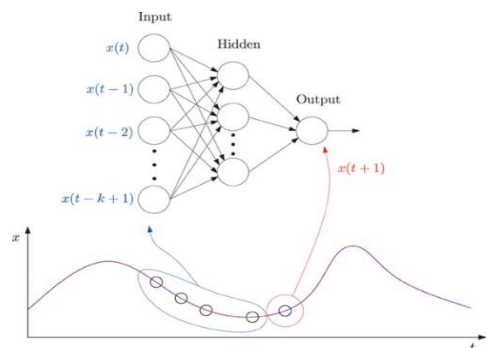
연구 과제 설명 등 홍보 내용 작성

시계열 데이터는 일정 시간 간격으로 배치된 데이터들의 수열을 말한다. 간단히 접할 수 있는 예로 주식 시장의 주가 변동 차트 등을 떠올려보면 된다.



구글 데이터 센터는 시간 별 서버들의 자원 요구량등의 정보를 담고있는 trace data를 대중에게 공개했다. 위의 그래프는 하루 동안의 서버의 CPU와 메모리 사용률의 변화를 보여준다.

시계열 예측을 정확하게 할 수 있다면 이 정보를 유용하게 사용할 수 있다. 주가의 변동을 미리 예측하면 큰 돈을 벌 수 있듯이, 변동하는 자원 (CPU, MEM) 요구량을 미리 알 수 있다면, 적시에 추가 자원을 할당하여 어플리케이션의 성능 저하를 예방하는 등 한정된 자원을 더 효율적으로 사용할 수 있다.



어떻게 예측할 수 있을까? 시계열 데이터를 예측하기 위한 많은 기법들이 존재한다. 아주 간단하게 “오늘 낮 1시의 기온은 작년과 동일할 것이다”와 같은 방법부터. 위 그림 처럼 Neural Network 를 이용하는 방법도 있다. 과제를 수행하면서 간단한 방법으로도 충분히 예측이 가능한지, 정확한 예측을 위해서는 복잡하고 많은 연산이 필요한지 등의 연구를 수행하게 된다.

과제의 목표는 시계열 예측 기법을 구글이 공개한 trace data 에 적용해 정확한 예측 및 시계열 생성의 법칙을 찾는 것이다.

과제를 수행하기 위해 익숙한 프로그래밍 언어 하나가 필요하다. 대부분 프로그래밍 언어로 시계열 데이터 예측을 위한 라이브러리가 존재하기 때문에 과제 수행에 큰 부담을 갖지 않아도 된다.



서울대학교 컴퓨터공학부
Seoul National University
Dept. of Computer Science and Engineering