

<https://doi.org/10.7236/JIIBC.2023.23.2.67>
JIIBC 2023-2-9

클라우드 기반 리눅스 SW 실습 플랫폼의 설계 및 구현

Design and Implementation of a Cloud-based Linux Software Practice Platform

반효경*, 조경운*

Hyokyung Bahn*, Kyungwoon Cho*

요약 최근 SW 실습 교육을 위해 수강생 별로 PC를 구비하는 대신 클라우드 상의 가상 PC를 할당해서 운영하는 사례가 늘고 있다. 본 논문에서는 클라우드를 이용해서 각 수강생의 SW 실습 환경을 효율적으로 구축할 수 있는 리눅스 기반 플랫폼을 설계하고 구현하였다. 본 논문의 플랫폼에서는 강사가 모든 수강생의 가상머신 템플릿을 일괄적으로 생성 및 통제할 수 있으며, 수강생들은 한 학기 동안 관리자 권한으로 본인의 머신을 관리하고 실습을 수행할 수 있다. 이때, 개별 머신에 발생하는 문제를 강사가 해결하거나 복원하는 효율적인 기능 또한 제공된다. 한편, 이러한 방식을 클라우드에서 실현하기 위한 가장 큰 어려움은 소요 비용을 미리 예측하기 어렵다는 데에 있다. 본 논문에서는 클라우드 이용 요금을 실시간으로 예측할 수 있는 모델을 설계하고, 사용자별 가상머신에 데몬을 두어 자원의 사용량 및 요금을 실시간으로 예측한다. 본 논문의 방식이 오버헤드가 거의 없음에도 실제 클라우드 사업자가 측정한 자원 사용량을 정확하게 예측하는 것을 확인할 수 있었다. 제안한 모델의 검증에 위해 리눅스 실습 강좌에서 한 학기 동안 클라우드를 활용한 결과 실제 청구 요금과 제안한 모델의 결과가 매우 유사한 것으로 확인되었다.

Abstract Recently, there are increasing cases of managing software labs by assigning virtual PCs in the cloud instead of physical PCs to each student. In this paper, we design and implement a Linux-based software practice platform that allows students to efficiently build their environments in the cloud. In our platform, instructors can create and control virtual machine templates for all students at once, and students practice on their own machines as administrators. Instructors can also troubleshoot each machine and restore its state. Meanwhile, the biggest obstacle to implementing this approach is the difficulty of predicting the costs of cloud services instantly. To cope with this situation, we propose a model that can estimate the cost of cloud resources used. By using daemons in each user's virtual machine, we instantly estimate resource usage and costs. Although our model has very low overhead, the predicted results are very close to the actual resource usage measured by cloud service providers. To further validate our model, we used the proposed platform in a Linux practice lecture for a semester and confirmed that the proposed model is very accurate.

Key Words : Cloud, Linux, PC virtualization, virtual machine, software practice

*정회원, 이화여자대학교 컴퓨터공학과
접수일자 2023년 3월 15일, 수정완료 2023년 3월 29일
게재확정일자 2023년 4월 7일

Received: 15 March, 2023 / Revised: 29 March, 2023 /
Accepted: 7 April, 2023

*Corresponding Author: bahn@ewha.ac.kr
Dept. of Computer Engineering, Ewha University, Korea

I. 서 론

최근 SW 실습실에 PC를 구비하는 대신 저사양 터미널을 설치하고 원격의 클라우드에 실습 환경을 생성해서 프로그래밍 및 시스템 관리 교육을 하는 사례가 늘고 있다. 특히, 리눅스 서버를 1대만 운영하고 수강생에게 해당 서버의 일반 사용자 계정을 발급하는 것이 아니라, 수강생마다 서버 관리자로서의 실습을 수행하기 위해서는 각 수강생에게 독립적인 머신을 할당하고 그 위에 리눅스가 설치된 형태의 시스템이 강의 기간 동안 지속적으로 유지되어야 한다. 이러한 방식은 실습실에 실제 하드웨어를 두지 않고도 모든 수강생이 고유의 하드웨어를 보유한 것과 같은 효과를 누릴 수 있다.

또한, 이러한 방식은 초기 장비 구입 없이 사용한 컴퓨팅 자원만큼만 비용을 지불하므로 하드웨어의 유지보수가 필요 없고 유휴 시간으로 인한 자원의 낭비도 없어 SW 실습 강좌에서 활용하기에 비용 효율성이 높다. 본 논문에서는 클라우드를 이용해서 각 수강생의 SW 실습 환경을 효율적으로 구축할 수 있는 리눅스 기반 플랫폼을 설계하고 구현하였다. 본 논문의 플랫폼에서는 강사가 모든 수강생의 가상머신 템플릿을 일괄적으로 생성 및 통제할 수 있으며, 수강생들은 한 학기 동안 관리자 권한으로 본인의 머신을 관리하고 실습을 수행할 수 있다. 이때, 개별 머신에 발생하는 문제를 강사가 해결하거나 복원하는 효율적인 기능 또한 제공된다.

그러나, 이러한 방식을 아마존 AWS와 같은 공용 클라우드에서 실현하기 위해 가장 큰 문제는 소요 비용을 미리 예측하기 어렵다는 데에 있다^[1, 2, 3]. 클라우드는 가상머신을 구성하는 사양에 따른 기본 단가와 해당 가상머신에서 실제 이용한 자원의 양에 의해 결정되는데 클라우드 운영자가 사용 자원에 따른 비용을 미리 알려주지 않기 때문에 적절한 대처를 하기까지 오랜 시간이 소요되기 때문이다.

본 논문에서는 SW 실습실에서 활용되는 클라우드 머신의 사용 요금을 미리 예측하고 이를 통해 효율적으로 클라우드 기반의 리눅스 SW 실습 강좌를 운영할 수 있는 방안에 대해 연구한다. 정확한 클라우드 이용 요금은 자원의 사용량을 측정해야 알 수 있으며, 이러한 자원은 CPU, 스토리지, 네트워크 등으로 구성된다. CPU의 경우 모델의 컴퓨팅 능력에 따른 기본 단가가 주어지고 이를 사용한 시간에 비례해 요금이 결정된다. 네트워크 자원의 경우 인바운드 및 아웃바운드 네트워크를 통한 데이터의 전송량을 측정하고 이에 근거해 요금이 결정된

다. 스토리지의 경우 저장 데이터의 용량과 I/O 접근 횟수를 동시에 고려하여 요금이 결정된다.

한편, 자원 사용량에 근거해 클라우드 요금이 결정되지만, 이는 클라우드 서비스 사업자가 측정해서 일방적으로 통보하는 방식이며 사용자는 청구된 요금에 대해 검증할 수 있는 방법이 없다^[4, 5, 6]. 따라서, SW 실습 강좌와 같이 다수의 사용자가 클라우드를 이용하는 경우 청구된 비용만큼 실제로 자원을 사용했는지 여부를 확인하기가 매우 어렵다.

본 논문이 제안하는 모델은 리눅스 기반의 SW 실습 강좌를 대상으로 공용 클라우드의 자원 사용량을 주기적으로 모니터링하여 요금을 실시간으로 예측할 수 있는 방법을 제시한다. 본 논문에서는 각 사용자 가상머신의 게스트 운영체제 상에서 구동되는 데몬을 통해 자원 이용 현황을 주기적으로 모니터링하고 이에 근거해서 자원 사용량 및 비용을 예측한다. 본 논문의 방식이 오버헤드가 거의 없음에도 실제 클라우드 사업자가 측정한 자원 사용량을 정확하게 예측하는 것을 확인할 수 있었다. 제안한 모델의 검증을 위해 리눅스 기반의 시스템 SW 실습 강좌를 한 학기 동안 운영하면서 AWS 공용 클라우드를 활용한 리눅스 실습을 수행한 결과 실제 클라우드 청구 요금과 모델에 의한 예측치가 매우 유사한 것으로 확인되었다.

II. 리눅스 기반 SW 실습 플랫폼 설계

다양한 운영체제와 툴이 필요한 SW 교육 과정을 지원하기 위해서는 여러 실습실이 별도로 구축되어야 하며, 동일한 플랫폼 하에서도 여러 강좌의 수강생이 한 대의 PC를 공유하는 형태가 되어 개인화된 관리에는 한계가 있다. 이를 해결하기 위해 멀티 부팅, PC 가상화 등 다양한 형태의 시도가 있어 왔고, 최근에는 개인 노트북을 지참해서 실습을 하는 형태가 오히려 주목받고 있는 상황이다.

본 논문에서는 공용 클라우드를 이용해서 각 수강생에게 독자적인 리눅스 머신을 한 학기 동안 제공하는 방식을 템플릿에 기반한 PC 가상화 기술로 해결하는 방안을 제안한다. 본 논문이 제안하는 SW 실습용 클라우드 솔루션에서는 각 수강생이 한 학기 동안 본인만의 컴퓨팅 환경이 설정된 리눅스 가상머신을 할당받게 되며, 강사는 해당 강좌에 적합한 템플릿으로 가상머신을 정의하여 세팅할 수 있도록 설계되었다. 또한, 세팅된 가상머신 템플

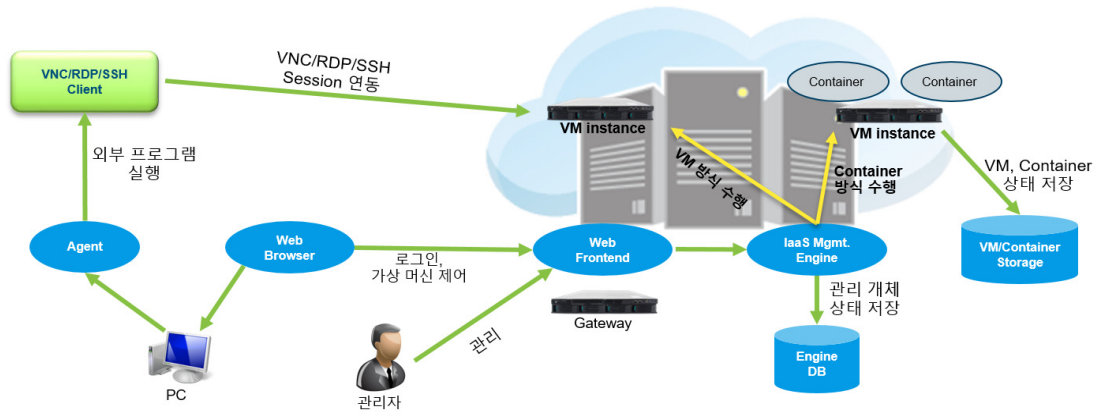


그림 1. 클라우드 기반 리눅스 SW 실습용 플랫폼의 구조
 Fig. 1. Architecture of a cloud-based Linux software practice platform.

릿이 각 수강생에게 생성 및 배포된 이후에도 강좌를 위한 라이브러리나 애플리케이션을 추가로 설치하고 이를 각 수강생의 가상머신에 미러링하는 작업이 용이하게 설계되었다. 한편, 각 수강생에게 관리자 권한이 부여되기 때문에 시스템의 중요한 설정이나 파일을 잘못 건드리거나 부팅 자체가 안 되는 등 시스템에 치명적인 상황이 발생할 수 있다. 본 논문에서 개발한 플랫폼은 이러한 상황에서 강사가 개별 학생들의 가상머신에 발생하는 문제를 해결하거나 시스템을 복원할 수 있게 하는 효율적인 기능 또한 제공한다.

수강생이 본인의 가상머신을 구동시킨 후 오랜 동안 사용하지 않을 경우 본 논문에서 개발된 솔루션은 불필요한 클라우드 비용의 발생을 막기 위해 자동으로 비활성화하는 방안을 제공하며, 각 수강생의 최대 사용 시간을 설정할 수 있도록 하여 과도한 비용 부담이 발생하지 않는 기능을 삽입하였다. 그림 1은 개발된 플랫폼의 구조를 보여주고 있다. 관리자는 수강생이 생성할 개체의 상태를 설정 및 관리하며 수강생은 각자의 PC에서 웹브라우저를 통해 클라우드의 상에 가상머신 혹은 컨테이너 인스턴스를 구동한다. 이때, 해당 인스턴스 상에 미터링 데몬이 실행되어 자원 사용량에 대한 모니터링이 이루어진다.

본 논문에서는 이러한 개발 솔루션을 바탕으로 한 학기 동안 실습 강좌를 운영할 경우 발생하는 클라우드 비용을 예측할 수 있는 모델을 제시하였다. 클라우드의 각 가상머신은 인스턴스로 불리며, 각 인스턴스의 상태는 활성 또는 비활성 상태로 분류할 수 있다. 활성 인스턴스의 비용은 CPU 비용, 스토리지 비용, 네트워크 비용으로 구분할 수 있으며, CPU 비용은 인스턴스 타입의 단가에 사용 시간이 곱해져서 결정된다. 스토리지 비용은 인스

턴스의 스토리지 볼륨 비용과 I/O 연산 비용이 합산되어 결정된다. 스토리지 볼륨 비용은 볼륨의 단가에 사용 시간이 곱해져서 결정되며, I/O 연산 비용은 단위 I/O 비용에 I/O 횟수 및 양을 고려한 누적값을 통해 결정된다. 네트워크 비용은 인바운드와 아웃바운드의 단위 네트워크 가격에 사용량을 곱해서 결정한다.

이때, 미리 결정되는 파라미터를 제외한 활성 인스턴스의 비용 예측을 실시간으로 하기 위해 본 논문에서는 주기적으로 스토리지 및 네트워크 이용량과 활성 인스턴스의 사용 시간을 모니터링한다. 비활성 인스턴스에서 발생하는 비용은 스토리지 볼륨의 비용이 유일하며, 이는 비활성 시간과 볼륨 이용 비용을 곱하여 예측할 수 있다.

III. 개발 플랫폼의 적용 및 검증

클라우드 사업자는 각 사용자 인스턴스의 자원 이용량을 토대로 요금을 결정한다. 그러나, 이러한 과정이 실시간으로 진행되지 않기 때문에 사용자는 즉각적인 비용 확인 및 이에 근거한 대처가 불가능하다^[7]. 본 논문이 제안하는 플랫폼에서는 각 사용자의 가상머신에 미터링 데몬을 두어 실시간으로 이용량을 추적하고 이에 근거해서 비용을 예측할 수 있도록 하였다. 비용 예측의 정확도는 아마존 AWS의 사용량 보고서를 통해 검증할 수 있다^[8]. 이를 위해 본 논문에서는 AWS 상에서의 다양한 인스턴스 및 사용 시나리오에 대해 제안 모델이 예측한 자원 사용량과 실제 AWS의 사용량 보고서를 비교하여 그 정확성을 검증한다. 구체적으로 리눅스 상의 다양한 애플리케이션을 포함하는 10종 이상의 시나리오로 인스턴스를

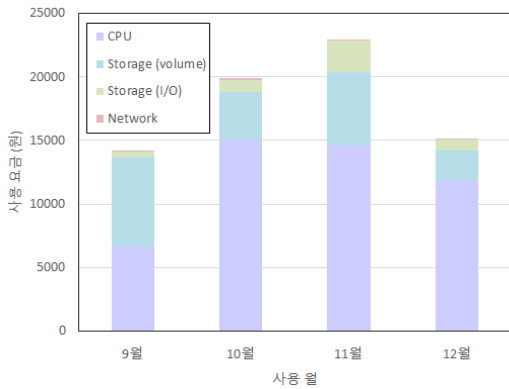


그림 2. 각 자원의 월별 비용 내역 분포
Fig. 2. Distributions of the cost for each resource type per month.

실행시켜 개발 플랫폼에 의한 비용 예측치와 AWS 보고서의 결과치를 비교하였다. 그 결과 본 논문의 플랫폼은 각 자원별 사용량 및 사용 시간을 99%까지 정확하게 예측할 수 있는 것으로 확인되었다.

이러한 검증 결과를 토대로 본 논문에서 개발한 솔루션을 이용하여 한 학기 동안 시스템 SW 강좌에서 AWS 클라우드 기반의 리눅스 SW 실습 플랫폼을 운영하였으며, 개발된 플랫폼에 의한 예측 비용과 실제 AWS에서 청구된 비용을 비교하였다. 그림 2는 AWS로부터 각 월별로 청구된 자원별 비용을 보여주고 있다. 그림에서 보는 것처럼 CPU 비용이 전체 기간에 걸쳐 가장 큰 비율을 차지했으며, 스토리지 볼륨의 비용은 월별 편차가 다소 큰 것을 확인할 수 있었다. 좀 더 상세한 분석 결과 이러한 스토리지 볼륨 비용의 대부분은 비활성 인스턴스에서 청구된 것으로 확인되었다. 네트워크 비용과 스토리지 I/O 비용은 모든 기간에 걸쳐 매우 작은 비율을 차지 하는 것을 확인할 수 있었다. 수강생용 가상머신들의 월별 총 비용은 25,000원 이내였으며 이는 AWS의 리눅스 인스턴스 중 매우 저렴한 사양을 사용했고 장시간 사용하지 않는 인스턴스를 자동 비활성화시키는 기능을 구현했기 때문으로 파악된다.

그림 3은 본 논문의 플랫폼을 통해 예측된 총 인스턴스의 비용과 실제 AWS가 청구한 요금을 시간이 흐름에 따라 비교해서 보여주고 있다. 그림에서 선그래프는 예측된 비용과 실제 청구 요금을 비교해서 보여주고 있으며 y축의 좌측에 구체적인 금액이 표시돼 있다. 그림에서 보는 것처럼 이 둘은 매우 유사한 값을 나타내었으며, 그 정확도는 99% 이상임을 확인할 수 있었다. 그림에서 막대그래프는 오차비용을 나타내며, y축의 우측에 금액을

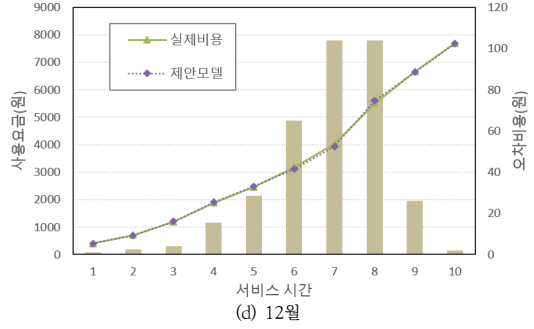
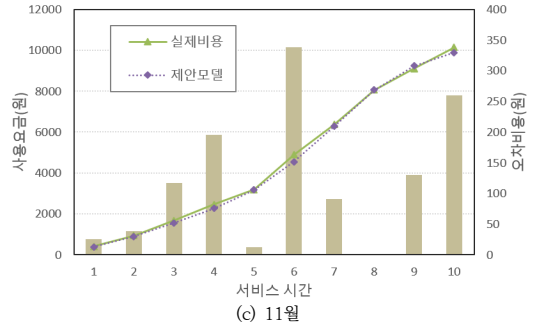
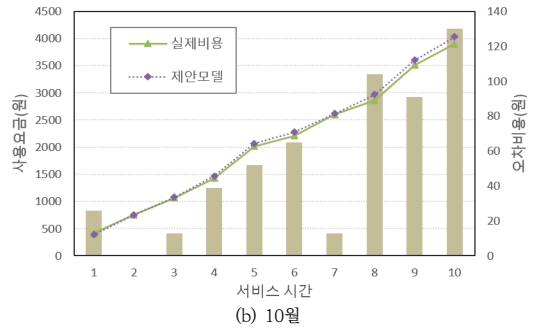
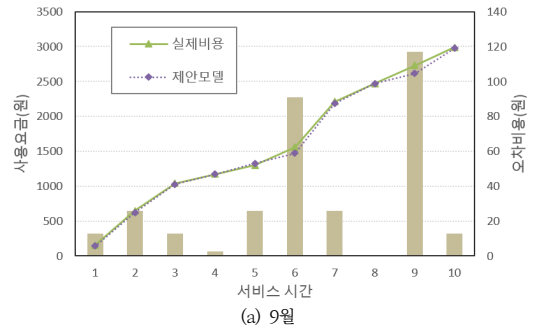


그림 3. 제안 모델과 실제 청구 비용의 비교
Fig. 3. Comparison of the proposed model and the actual cost.

표시해서 보여주고 있다. 비록 일부 구간에서 오차가 다소 높아지지만 월 단위로 볼 때 항상 3% 미만의 오차가 발생하며 평균 1% 이내임을 확인할 수 있었다.

IV. 결 론

SW 실습실에 PC를 구비하는 대신 클라우드 상의 가상 머신을 통해 실습을 수행하는 방식이 각광받고 있다. 이러한 방식은 관리의 효율성뿐 아니라 개인화된 플랫폼을 한 학기 동안 유지할 수 있어 낮은 비용으로 모든 수강생에게 물리적인 머신을 1대씩 제공할 것과 유사한 효과를 얻을 수 있다. 그러나, 공용 클라우드의 이용 요금을 미리 예측할 수 없기 때문에 이러한 방식을 사용하는 데에는 어려움이 있다. 본 논문에서는 강사가 일괄적으로 모든 수강생에게 가상머신 템플릿을 생성 및 배포하고 수강생은 한 학기 동안 해당 가상머신을 본인만의 고유한 머신으로 사용할 수 있는 클라우드 기반 리눅스 SW 실습 플랫폼을 개발하였다. 또한, 본 논문에서는 수강생의 가상머신에서 소요되는 클라우드 비용을 실시간으로 파악할 수 있는 모델을 개발하고 장시간 사용하지 않는 가상머신을 비활성화하는 기능을 탑재하였다. 본 논문의 플랫폼이 공용 클라우드의 자원 사용량을 정확하게 예측하는 것을 실제 클라우드 사업자가 측정할 결과와의 비교를 통해 검증하였으며, 개발된 플랫폼으로 SW 실습 강좌를 한 학기 동안 운영하여 공용 클라우드에서 활용 가능한 방법임을 입증하였다.

References

- [1] M. Al-Roomi, S. Al-Ebrahim, S. Buqrais, and I. Ahmad, "Cloud computing pricing models: a survey," *International Journal of Grid and Distributed Computing*, vol. 6, no. 5, pp. 93-106, 2013.
DOI: <https://doi.org/10.14257/ijgdc.2013.6.5.09>
- [2] S. Yoon and Y. Lee, "Cloud based distributed parallel processing for deep learning based real-time person detection in an AI middleware for smart city," *The Journal of KIIT*, vol. 20, no. 5, pp. 9-19, 2022.
DOI: <https://doi.org/10.14801/jkiit.2022.20.5.9>
- [3] J. Park and E. Park, "Performance evaluation of IoT cloud platforms for smart buildings," *Journal of the Korea Academia-Industrial cooperation Society(JKAIS)*, vol. 21, no. 5, pp. 664-671, 2020.
DOI: <https://doi.org/10.5762/KAIS.2020.21.5.664>
- [4] A. Anwar, A. Sailer, A. Kochut, C. O. Schulz, A. Segal,

and A. R. Butt, "Scalable metering for an affordable IT cloud service management," in *Proc. IEEE International Conference on Cloud Engineering*, pp. 207-212, 2015.
DOI: <https://doi.org/10.1109/IC2E.2015.18>

- [5] M. Artan, I. Shabani, and B. Sejdiu, "Pricing schemes in cloud computing: an overview," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 2, pp. 80-86, 2016.
- [6] P. Mitropoulou, E. Filiopoulou, M. Nikolaidou, and C. Michalakelis, "Pricing IaaS: a hedonic price index approach," *Lecture Notes in Computer Science*, vol. 10537, pp. 18-28, 2017.
DOI: https://doi.org/10.1007/978-3-319-68066-8_2
- [7] B. Martens, M. Walterbusch, and F. Teuteberg, "Costing of cloud computing services: A total cost of ownership approach," in *Proc. 45th IEEE Hawaii International Conference on System Sciences*, pp. 1563-1572, 2012.
DOI: <https://doi.org/10.1109/HICSS.2012.186>.
- [8] AWS Cost and Usage Reports,
<https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/billing-reports-costusage.html>.

저 자 소 개

반 호 경(정회원)



- 1997년 2월 : 서울대학교 계산통계학과 학사
- 1999년 2월 : 서울대학교 전산과학과 석사
- 2002년 2월 : 서울대학교 컴퓨터공학부 박사.
- 2002년 9월 ~ : 이화여자대학교 컴퓨터공학과 교수.

• 주관심분야 : 운영체제, 스토리지시스템, 임베디드시스템

조 경 운(정회원)



- 1995년 2월 : 서울대학교 계산통계학과 학사
- 1997년 2월 : 서울대학교 전산과학과 석사
- 2012년 2월 : 서울대학교 컴퓨터공학부 박사
- 2000년 ~ 2016년 : (주)클루닉스 연구소장

• 2016년 4월 ~ : 이화여자대학교 임베디드소프트웨어연구센터 수석연구원/연구교수

※ This work was partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1A2C1009275) and the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2021-0-02068, Artificial Intelligence Innovation Hub).