Neural Network를 이용한 강화학습 기반의 잡샵 스케쥴링 접근법

정현석[O*], 김민우[*], 이병준[*], 김경태[*], 윤희용[**]
[O*]성균관대학교 정보통신대학 전자전기컴퓨터공학과
[**]성균관대학교 소프트웨어대학 소프트웨어학과
e-mail: {daun8928, kimmw95, byungjun}@skku.edu[O*], kyungtaekim76@gmail.com[*], youn7147@skku.edu[**]

An Neural Network Approach to Job-shop Scheduling based on Reinforcement Learning

Hyun-Seok Jeong[O*], Min-Woo Kim[*], Byung-Jun Lee[*], Kyoung-Tae Kim[*], Hee-Yong Youn[**]
[O*]Dept. of Electrical and Computer Engineering, Sungkyunkwan University
[**]Dept. of Software, Sungkyunkwan University

● 요 약 ●

본 논문에서는 NP-hard 문제로 알려진 잡샵 스케쥴링에 대하여 강화학습적 측면에서 접근하는 방식에 대해 제안한다. 다양한 시간이 소요되는 업무들이 가지는 특징들을 최대한 state space aggregation에 고려하고, 이를 neural network를 통해 최적화 시간을 줄이는 방식이다. 잡샵 스케쥴링에 대한 솔루션은 미래에 대한 예측이 불가능하고 다양한 시간이 소요되는 스케쥴링 문제를 최적화하는 것에 대한 가능성을 제시할 것으로 기대된다.

키워드: 강화학습(reinforcement learning), 잡샵 스케쥴링(job-shop scheduling), IIoT(industrial internet of things), 신경망(neural network)

I. Introduction

IIoT 기술의 발전에 따라 공정은 방대한 센서 데이터와 컴퓨팅 파워를 갖게 될 것이다. 대부분의 업무들은 한정된 자원을 분배해서 써야 하는 것이며, 방대한 데이터만큼 스케쥴링의 중요도가 올라간다. 패킷 스케쥴링부터 공정 자체의 처리 순서까지 매우 다양한 범위에서 스케쥴링이 요구된다.

잡샵 스케쥴링에 관한 일반적인 해법은 알려져 있지 않다. 일반적으로 domain-specific 한 범위로 제한하는 이유도 이것이다. 본 논문에서는 IIoT 환경에서 여러 가지 CPU 자원, 다양한 길이의 패킷과, 다양한 패킷 처리시간, 다양한 처리 요구 순서가 존재하는 환경에서의 스케쥴링 방식을 제안하고자 한다. 기존 관련 연구는 휴리스틱과 제한시간을 통하여 접근한다. 본 논문에서는 이와 같은 접근을 받아들이고, 컴퓨팅 파워에 의존할 수 있는 강화학습 기반의 접근에 대한 가능성을 발견해보고자 한다.

본 논문은 1장 소개를 다음으로, 2장에서는 관련연구, 3장에서는 제안하는 기법, 4장에서는 결론으로 구성하였다.

II. Preliminaries

1. Related works

1.1 Job-Shop Scheduling

스케쥴링의 목적은 할당된 업무를 제한된 리소스에 최적화된 방식으로 분배하는 것이다. 잡샵 스케쥴링에선 n개의 job이 m개의 machine에서 특정한 순서로 처리되어야 한다. 각 job을 j라고 표시하면 v_j개의 operation으로 이루어지고, 이는 $o_{j,v_j},\ldots,o_{j,v_j}$와 같은 식으로 표현한다. 이들은 정해진 시간동안 특정한 리소스를 사용하며, 모든 operation을 끝마쳐야 job이 완료된다.

1.2 강화학습

다음으로 강화학습을 통해 가장 높은 보상이 기대되는 선택하는 알고리즘이 제시되어 있다. 환경으로부터의 상호작용을 통해 보상을 제시하고, 현재의 상태로부터 얻어지는 보상을 기록한다. 이를 반복하여 최적의 해를 발견한다.

1.3 Multi Layered Neural Network

다수의 Layer가 네트워크로 연결되어 있고, 각 Layer는 순서를

가지고 있다. 각 Layer는 노드들을 포함하고 있는데, 앞선 Layer의 노드는 다음 Layer의 노드와 모두 연결되어있으며, 각 연결부위는 가중치를 가지게 된다. 각 노드는 또한 자신의 값을 가지고 있으며, 이전 Layer의 연결과 bias를 통해 다음 노드의 값을 정하게 된다.

III. The Proposed Scheme

잡샵 스케쥴링을 MDP 모델링 하는 데에 있어서는 일반적으로 두가지 방식이 있다. 하나는 스케쥴링을 single MDP 문제로 모델링 하는 방식이다. 또 다른 방식은 multi-agent MDP 문제로 모델링 하는 방식이다. 제안하는 방식으로는 multi-agent 기반의 MDP 문제로 모델링 하는 것이다. 이는 가능한 Action들을 몇 개의 element로 만든 뒤, 사용 가능한 리소스의 개수만큼 묶어서 하나의 벡터로 만든다.

스케쥴링 방식에는 예측적(predictive) 스케쥴링과 반응적(reactive) 스케쥴링이 있는데, multi-agent MDP에서는 반응적인 스케쥴링을 이용한다. 이는 완벽하게 예측 불가능한 미래의 사건들에 대해서도 적절히 대처할 수 있다는 장점이 있다.

현재 시간 t에 대해 모든 자원과 처리중인 작업에 관한 모든 특징을 포함하는 종합적인 view를 갖는 state $s(t)$는 최적화 가능성에 관해서는 매우 긍정적이다. 하지만 이는 많은 state space를 필요로 하고, 일반적으로는 적절하게 추려져야만 실시간으로 적용 가능하다.

State space를 지나치게 추리지 않으면서도 이를 해결하기 위해서 Multilayer perceptron neural network가 state-action function을 대표할 수 있도록 한다. neural network는 어떤 함수든 최대한 근사하여 표현하는 것이 가능한 것으로 알려져있다. 또한 Feature Learning을 하여 해당 함수의 특징을 추출해낼 수 있다. 이는 방대한 Q-table을 전부 다 채우기 전에 적절한 수준으로 합리적인 해결책을 제시할 수 있으며, 방대한 데이터 처리량 때문에 스케쥴링 중요도가 높은 곳에서 현실적으로 적용 가능할 것으로 보인다.

IV. Conclusions

본 논문에서는 다양한 길이의 job과 다양한 요구 조건을 지닌 잡샵 스케쥴링 하기 위해 neural network를 이용하는 방식을 제안하였다. 대부분의 패킷 스케쥴링은 리소스 할당량의 비율을 정하거나, 모든 패킷의 처리량을 동일하게 가정한 뒤 우선순위 큐 혹은 가중치를 이용해서 해결을 하고자 한다. 이는 실시간 시스템이 가지는 제약 때문이기도 하다. 하지만 본 논문에서는 이외의 특징들을 모두 고려할 수 있는 스케쥴링 방식을 제안했다. 현실적으로 적용 가능한 수준의 해결책이 제시될 수 있도록 추가 연구가 이루어질 예정이다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신·방송연구 개발 사업(No. 2016-0-00133, 초연결 IoT 노드의 군집 지능화를 통한 Edge Computing 핵심 기술 연구), SW중심대학지원사업(2015-0-00914), 한국연구재단 기초연구사업(No.2016R1A6A3A11931385, 실시간 공공안전 서비스를 위한 소프트웨어 정의 무선 센서 네트워크 핵심기술 연구, 2017R1A2B2009095, 실시간 스트림 데이터 처리 및 Multi-connectivity를 지원하는 SDN 기반 WSN 핵심 기술 연구), 삼성전자, BK21PLUS 사업의 일환으로 수행되었음.

REFERENCES

[1] R. Suryaputra, and Williem, "Reinforcement Learning Combined with Radial Basis Function Neural Network to Solve Job-Shop Scheduling Problem," IEEE International Summer Conference of Asia Pacific Business Innovation and Technology Management, pp. 29-32, July 2011.

[2] Y. Wei, and M. Zhao, "Composite rules selection using reinforcement learning for dynamic job-shop scheduling," IEEE Conference on Robotics, Automation and Mechatronics, Vol. 2, pp. 1083-1088 Dec. 2004.