J. Parallel Distrib. Comput. 75 (2015) 93–100

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

journal homepage: www.elsevier.com/locate/jpdc

# Improving reliability in resource management through adaptive reinforcement learning for distributed systems

Masnida Hussin [a,*], Nor Asilah Wati Abdul Hamid [a], Khairul Azhar Kasmiran [b]

[a] Department of Communication Technology & Networks, Faculty of Computer Science & IT, 43400 University Putra Malaysia (UPM), Selangor, Malaysia

[b] Department of Computer Science, Faculty of Computer Science & IT, 43400 University Putra Malaysia (UPM), Selangor, Malaysia

## HIGHLIGHTS

- Able to drive the evolution of automation networks towards higher reliability.
- Able to handle tasks at different states in processing.
- Able to adapt with system changes while leading better learning experiences.
- Able to sustain reliable performance
- Able to deal with dynamic global network infrastructure.

## ARTICLE INFO

*Article history:*
Received 10 January 2014
Received in revised form
13 September 2014
Accepted 2 October 2014
Available online 13 October 2014

*Keywords:*
Distributed systems
Resource management
Adaptive reinforcement learning
System reliability
Computational complexity

## ABSTRACT

Demands on capacity of distributed systems (e.g., Grid and Cloud) play a crucial role in today's information era due to the growing scale of the systems. While the distributed systems provide a vast amount of computing power their reliability is often hard to be guaranteed. This paper presents effective resource management using adaptive reinforcement learning (RL) that focuses on improving successful execution with low computational complexity. The approach uses an emerging methodology of RL in conjunction with neural network to help a scheduler that effectively observes and adapts to dynamic changes in execution environments. The observation of environment at various learning stages that normalize by resource-aware availability and feedback-based scheduling significantly brings the environments closer to the optimal solutions. Our approach also solves a high computational complexity in RL system through on-demand information sharing. Results from our extensive simulations demonstrate the effectiveness of adaptive RL for improving system reliability.



## 1. Introduction

Large-scale distributed systems including Grids and Clouds are driven by an expansion of the Internet that able to provide massive information and dynamic computing services. Heterogeneity and dynamicity of resources and applications in these systems are rather common and must be effectively dealt with [11]. Resource allocation considering these heterogeneous and dynamic characteristics has become increasingly more important with the advent and prevalence of the distributed systems, e.g., Amazon Elastic Compute Cloud (EC2) as a public cloud. However, the diverse nature of network devices/components and communication technologies highly increases complexity in resource management. It poses a number of new technical challenges including performance variability and accountability. The demand of system reliability for massive and complex networked environment imposes much burden on resource allocation solutions. The changing nature of network services and a large number of devices connected in distributed systems demand for new resource allocation approaches to efficient use of heterogeneous resources.

Robust and reliable network services depend heavily on the quality of resource allocation (scheduling) decisions for improving overall network performance [6,12]. The decisions are not only taking charge of matching and scheduling the processing capacities of resources and user requirements, but they also need to deal with a wide variety of resource behaviors and performance fluctuations. Services in large-scale distributed environments (e.g., The Internet of Things [15]) are required to stay and continue operating even in the presence of malicious and unpredictable

* Corresponding author.
*E-mail addresses:* masnida@upm.edu.my (M. Hussin), asilah@upm.edu.my (N. Asilah Wati Abdul Hamid), k_azhar@upm.edu.my (K.A. Kasmiran).

http://dx.doi.org/10.1016/j.jpdc.2014.10.001
0743-7315/

circumstances; that is, their processing capacity must not be significantly affected by the user requirements [11,8]. For instance, the scheduling decisions are required to be reliable in the sense of tolerating failure automatically, or guaranteeing properties such as high availability; hence having good performance even under malfunctioned. However, it raises complexity to adjust the resource allocation decisions in highly dynamic and complex network environment where there is a need for making quick solutions without significant performance degradation.

We address adaptive and scalable resource allocation in the face of uncertainty. These abilities are important for trustworthy scheduling decisions. Mainly, the distributed systems require consistent and iterative monitoring for valuation resources' behaviors and processing requirements. Therefore, an autonomous, scalable and highly dynamic learning approach is deserved. Recently, a promising approach based on reinforcement learning (RL) has been studied for dynamic task scheduling and resource allocation [18,16]. RL offers an effective decision-making policy while providing with less system-specific knowledge that can be of a very practical scheme for resource allocation and scheduling. It is also a promising approach for automatically developing effective and efficient policies for real-time self-adaptive management. Since the RL system is capable of repairing and improving its performance when exhibiting poor in the quality of decisions [16], the dependable information for future decision-making is realized. However, it raises computational complexity when the problem space is too large to explore completely, hence it is hard to predict about new situations and to find an optimal solution. The reinforcement-agent is required to effectively explore the environment while modifying its action in response to system changes such as performance fluctuations, high computing complexity and massive distributed applications. An online and fast-adaptive learning to fabricate sufficient knowledge for reliable experiences is necessary that helps the agent to sense, manipulate and address information at once.

In this paper, a novel **D**ynamic-**A**daptive resource allocation based on **RE**inforcement **L**earning (*DAREL*) is proposed and evaluated. There are many different ways (e.g., [18,16,14,17]) to incorporate reinforcement learning in resource allocation. In this work, we combined an Artificial Neural Network (ANN) [18] with stochastic automata for handling environments that are not deterministic. Our model-based reinforcement learning strategy aims to simultaneously learn system environment and improve learning policy. It realizes by performing actions (or allocation decisions) and observing feedbacks while memorizing the actions that bring the systems closer to the optimal decisions. The reinforcement-agents autonomously explore and respond to its environment for allocating applications into resources based on rewarding system. In order to adapt with the system changes and scale of learning experiences, *DAREL* also supports an *on-demand* information sharing strategy that significantly accelerates a convergence speed of learning. Our experimental results demonstrate that *DAREL* plays a significant role in enhancing system reliability, also reduces computational complexity in RL.

The reminder of this paper is organized as follows. A review of related work is presented in Section 2. In Section 3, we describe the models used in the paper. Section 4 presents the structure of reinforcement learning (RL) for dynamic and massive network environment. In Section 5, we detail the components of *DAREL* for reliable services. Experimental setting, comparison of different resource scheduling policies and results are presented in Section 6. Finally, conclusions are made in Section 7.

## 2. Related work

Effective resource allocation mainly consider some substantial issues such as effective matching and scheduling strategies, usage status of resources, communication between schedulers and approximate task execution time. Developing self-adaptive scheduling policy to deal with variability of application services is a key issue in optimizing utility distributed systems. One of the strategies that is proposed for resource allocation problems is reinforcement learning approach [18,16,10,2]. In [18], the authors introduced RL-based dynamic scheduling policy in parallel processors systems. The work highlights scalability and adaptability as the main key objectives of dynamic scheduling in decentralized learning strategy. It also shows the effectiveness of RL-based scheduler model towards improving the quality of scheduling decisions, through on-going learning process.

The authors in [16] proposed the method called ordinal sharing learning (*OSL*) to cope with exploration and exploitation abilities of the agents. For increasing the learning exploration, the agents interact with the environment using an indirect method where they use the applications' information to estimate the state. Meanwhile, the utility table that is updated by each agent is ordinal and iteratively shared in neighborhood to view and estimate the efficiency of resources. As such, efficient coordination among agents and optimum utilization in the proposed learning strategy are realized. However, they do not formulate the decisions for highly dynamic environments where uncertainties and variability of distributed resources and application are present constantly. Our work adopts the distributed reinforcement learning structure and learning-feedback loop in [18,16] for better resource scheduling that focuses for highly dynamic and probabilistic environments.

The learning-based scheduling policy mainly aims to build up knowledge for improving the ability of the scheduler in solving uncertain decision-making problem. The control learning strategy that is proposed in [2] develops a policy to identify the most appropriate action from different states. Their reinforcement-agents learn by trial and error to identify the optimal action for a given state. Our work focused more on reinforcement learning strategy for resource allocation compared to agent capability that is highlighted in [2]. The adaptive reinforcement learning in [1] improves online learning policy for Web systems auto-configuration. However, they do not formulate RL framework especially for initial learning phase of training data. For optimal allocation decisions, the study in [9] proposed negotiation-based resource allocation using the $Q$-learning technique (*NQL*). The system compositions (users and providers) make successive offers that depend on very quick negotiation process. It identifies the possible negotiation state and selects the action with the highest learning rate. The agent of each resource provider behaves greedily most of the time.

*QIA* in [10] sustains incentive in every resource in order to predict how well a provider can adjust its trust factor optimally. The trust factor that is computed through the Markov Decision Processes (MDP) method aims to recognize dependable computing node among available resources. During the decision-making process, the agent learned the observed the trust factor, the resource with highest $Q$-value is then chosen. The learning strategy that is based on ordinal information sharing [16], negotiation [9] and incentive [10] for resource allocation serves as a point of comparison in our experimental evaluation. Their strategies address average distributed system utility in the learning policy that implies in our work as well.

## 3. Model

In this section, we describe the system and application models employed in this work.

### 3.1. System model

The target network system used in this work consists of multiple resource sites that are loosely connected by a communication