# Reinforcement Learning in the domain of Resource Optimization

Reinforcement Learning Assignment for IE7860: Intelligent Analytics, WAYNE STATE UNIVERSITY, Dept. of Industrial & Systems Engineering, Winter 2020

Review of several papers describing application of Reinforcement Learning techniques to the problem domain of Resource Management, in particular Computational Resources.

By Igor Ostapthenko AKA igor\_ost@wayne.edu

## Application of traditional Machine Learning to the Resource Management Domain

The resource management is “.. the efficient and effective development of an organization’s resources when they are needed.” according to Wikipedia. The resources may include the financial resources, inventory, a set of weapon systems with limited ammo, combat capabilities and survivability, or natural resources. I was always interested in the general theory of resource management, especially in the realm of computer science, where there is no such thing as unlimited storage or bandwidth. Effective scheduling of jobs in the shared environment is hard problem, especially when environment is distributed. There are two major aspects of this domain I have studied. One is when the environment comprised of several standard components like web server, application servers and database, sometimes combined with the upper limit of the cpu, disk and bandwidth available due to space or other physical limitations. For this constraints one wants to find optimal set of tuning parameters for the changing computational demands over given time cycle: days of weeks, times of day. The example of these kind of problem being solved with the application of Machine Learning is the subject of the paper (Yi and Connan, 2007), describing KernTune. The idea of having a few pre-defined sets of parameters corresponding the the Class of Load, examples one being serving bid-requests over HTTP during the day and the the other is running Spark/Hadoop jobs as part of the YARN or DS/OS Cluster cluster during the night, will require different kernel, file system and cache parameters as well as alternative CPU/IO scheduling strategies. The clever application of the Support Vector Machines based classifier allowed to learn to map the observed system metrics the state system needs to be in. This is approach uses supervised learning and classification. Although there always be a delay in reaction the new set of tuning parameters will be deployed and various system components restarted or reloaded for new parameters to be used to match the new load pattern. For example deploying the applications to cluster different classes of load passing through the cluster can shift the parameters towards the optimal settings for the given load.

## Reinforcement Learning applications

The modern DataCenter going towards increasing system scale and complexity that introduce more and more configurable parameters to a level beyond the capacity of an average-skilled operator. The studies reveal that more than 50% root causes of Internet services outage are due to misconfigurations by the operator (Bu, Rao, Xu. 2009). Further in the article the authors offer the Reinforcement Learning approach to tune the most important parameters of Web server and Application server, the two workhorses of modern web applications, in VM-based dynamic environment. The proposed algorithm explores system dynamic features by interacting with the external environment. Th proposed methodology of online system configuration adaptation in the presence of variations and able to change the system to a near-optimal configuration within less than 25 iterations. The next level of application is considered in (Mao, Alizadeh, Menache, Kandula 2016). TheDeepRM introduced in the article deals with resources of the cluster not the single machine or VMs in the single machine. The Cluster Scheduling is helped by Reinforcement Learning algorithm, although the model is simplified: the cluster is a single collection of resources with no machine fragmentation and the scheduling is not pre-emptive. The DeepRM is compared to Shortest Job First (SJF) that allocates jobs in increasing order of their duration; a Packer agent which allocates jobs in increasing order of alignment between job demands and resource availability and Tetris agent that balances preferences for short jobs and resource packing in a combined score. The DeepRM agents are sometimes better than heuristics for a multi-resource cluster scheduling problem. Learning resource management strategies directly from experience, could offer a real alternative to current heuristic based approaches.

## Conclusion

Further research in this area will allow to move the knowledge about how to run massive computation at scale from the realm of wizards and rule-of-thumb towards the observed and learned understandings of the relationships in the system.

## References

* https://towardsdatascience.com/applications-of-reinforcement-learning-in-real-world-1a94955bcd12
* Yi, Long & Connan, James. (2007). KernTune: Self-tuning Linux kernel performance using support vector machines. 10.1145/1292491.1292513.
* Xiangping Bu, Jia Rao, Cheng-Zhong Xu. (2009): *A Reinforcement Learning Approach to Online Web Systems Auto-configuration* Department of Electrical & Computer Engineering Wayne State University, Detroit, Michigan 48202 {xpbu,jrao,czxu}@wayne.edu, 2009 29th IEEE International Conference on Distributed Computing Systems, DOI 10.1109/ICDCS.2009.76
* Hongzi Mao, Mohammad Alizadeh, Ishai Menache, Srikanth Kandula (2016) *Resource Management with Deep Reinforcement Learning* † Massachusetts Institute of Technology⋆, Microsoft Research† {hongzi, alizadeh}@mit.edu, {ishai, srikanth}@microsoft.com, https://dl.acm.org/doi/10.1145/3005745.3005750