Autonomic Management of Large-Scale Critical Infrastructures

Ivana Dusparic and Vinny Cahill
Distributed Systems Group, Department of Computer Science
Trinity College Dublin

1. Introduction to Large-Scale Critical Infrastructures

Critical Infrastructures include facilities, services and installations essential for the functioning of a society and economy. Such infrastructures are generally of a very large scale spanning cities, whole countries, or even crossing international borders. Examples of such Large-Scale Critical Infrastructures (LSCI) are electricity, water and gas supply, transportation, and health service [8]. There are many aspects to the management of LSCI, including security, fault tolerance, availability, and reconfiguration. The particular aspect of LSCI management with which we are concerned and which we view as a grand challenges in autonomic computing is optimization of their performance in changing conditions. Not all circumstances in which these systems will operate can be predicted so it is not possible to fully define their behaviour at design time. Even for the known operating conditions, with hundreds or thousands of nodes, it is infeasible, if not impossible, to define correct behaviour for all combinations of conditions on all nodes. Critical infrastructures need to adapt to various changes in load, both sudden ones and repeated load patterns. They need to optimize their performance with respect to multiple, often conflicting or highly dependent, policies with different levels of priority (high, low), affecting different parts of the systems (local, regional, global), either continuously, or in certain cirumstances (sporadically). In our work, Urban Traffic Control (UTC) is used as an exemplar of a LSCI. A UTC systems consist of hundreds of dependent nodes (traffic lights) that need to coordinate their behaviour to deal with optimizing general traffic throughput, prioritizing emergency vehicles and public transport, as well as adapting to any surges of traffic in particular areas in case of public events or accidents.

2. Self-Organizing Decentralized Management of LSCI

Due to the size of LSCI and the heterogeneity of their nodes and policies, we hypothesize that the most feasible approach to their management is self-management achieved through decentralized learning and cooperation. The global view of the system needed for centralized management is not possible as it would need to combine views of all nodes in the system and dictate their performance. In the fast changing circumstances in which these systems operate, such a global view has the potential of becoming outdated by the time it is compiled. Therefore, global consensus on performance should be achieved through cooperation between agents having only a local view of the system. Adaptation should be achieved through local processes of learning and feedback from neighboring agents. Techniques achieving such adaptation based on self-organizing biological systems and machine learning are already being used in management of large-scale distributed systems. For example, ant colony optimization is being applied in load balancing [7], particle swarm optimization in wireless networks [4] , evolutionary computing in routing [5] and traffic [6], and Reinforcement Learning (RL) [9] in dynamic resource allocation [10]. These techniques in their current implementations mostly focus on adapting performance w.r.t. a single system policy. However, one of the main requirements for self-organizing techniques to be used in management of LSCI is the ability to handle multiple policies simultaneously, as well as achieving global behaviour through the local behaviour of multiple individual agents and their cooperation. Modifications of RL exist that coordinate multiple agents' performance w.r.t. a single policy, as well as RL techniques that optimize for multiple policies on a single agent but no RL techniques combine both. We propose adapting existing RL techniques to implement multiple policies by coordinating behaviour of multiple agents without central control, in order to optimize performance of LSCI.

3. Multiple Policy Collaborative Reinforcement Learning

In our work, we are combining single-policy multi-agent RL technique known as Collaborative Reinforcement Learning (CRL)[2] and multi-policy single-agent RL techniques (W-learning [3] and combined state spaces [1]), and adapting them to the unique characteristics of LSCI. In LSCI, a high dependency between agents is present (e.g. in UTC actions at one junction directly influence the state of downstream junctions), as well as high dependency between policies (e.g. in UTC if we do not optimize the traffic flow, emergency vehicles will be delayed as well due to traffic jams).

3.1. First Results

Our first results are obtained from experiments on a single agent (junction) and two policies: minimizing global traffic waiting times, and prioritizing emergency vehicles. We tested two approaches to optimizing both policies: separate learning processes for each policy, where policies are competing for the control of the traffic light using W-learning, and a single learning process where a single state space encodes information relevant for both policies [Figure 1]. We compared both policies to a round-robin traffic light controller.

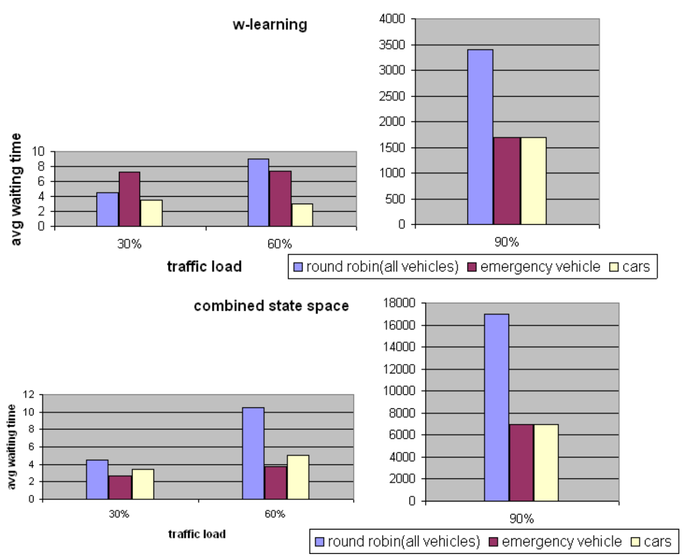

Figure 1. RL vs Round-Robin

Our first results show that for our chosen combination of policies (high priority, regional, sporadic policy prioritizing emergency vehicles vs low priority, global, continuous policy minimizing global vehicle waiting time) the combined state space shows better improvement over round robin, and clearly prioritizes the higher priority policy for all system loads. In W-learning policies are competing for control over the agent, so W-learning, even though it shows improvement over round robin, might be more suitable for policies of equal priority and equal frequency.

4. Future Work, Summary and Conclusions

Based on our first results, we see that suitable technique for optimization of multiple policies in LSCI highly depends on policy types and the relationship between the policies. Our goal is to conduct experiments on multiple agents in various relationships (e.g. independent junctions vs upstream/downstream relationship) and with various combinations of policies of different types and relationships (e.g. local vs global policy, high vs low priority), and establish patterns to derive a set of conclusions on what RL technique is suitable for which combination. Our initial conclusion is that self-organizing algorithms represent a promising solution for decentralized autonomic management of LSCI, but that different algorithms or different versions of the algorithms are suitable for different characteristics of the policies, agents, and system as a whole.

References

[1] H. Cuayáhuitl, S. Renals, O. Lemon, and H. Shimodaira. Learning multi-goal dialogue strategies using reinforcement learning with reduced state-action spaces. In *Proc. of Interspeech-ICSLP 2006*, 2006.

[2] J. Dowling. *The Decentralised Coordination of Self-Adaptive Components for Autonomic Distributed Systems*. PhD thesis, Trinity College Dublin, 2005.

[3] M. Humphrys. *Action Selection methods using Reinforcement Learning*. PhD thesis, University of Cambridge, 1997.

[4] B. A. Kadrovach and G. B. Lamont. A particle swarm model for swarm-based networked sensor systems. In *SAC*, pages 918–924, 2002.

[5] H. L. and M. N. Hybrid genetic algorithms for telecommunications network back-up routeing. *BT Technology Journal*, 2000.

[6] S. Mikami and Y. Kakazu. Genetic reinforcement learning for cooperative traffic signal control. In *International Conference on Evolutionary Computation*, pages 223–228, 1994.

[7] A. Montresor, H. Meling, and O. Baboglu. Messor: Load-balancing through a swarm of autonomous agents. In G. Moro and M. Koubarakis, editors, *Proceedings of the 1st Workshop on Agent and Peer-to-Peer Systems (AP2PC'02)*, number 2530 in Lecture Notes in Artificial Intelligence, pages 125–137, Bologna, Italy, 2003. Springer-Verlag.

[8] J. Moteff and P. Parfomak. Critical infrastructure and key assets:definition and identification. Technical report, Congressional Research Service, The Library of Congress, 2004.

[9] R. S. Suton and A. G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book. The MIT Press, Cambridge, Massachusetts, 2002.

[10] G. Tesauro. Reinforcement learning in autonomic computing: A manifesto and case studies. *IEEE Internet Computing*, 11(1):22–30, 2007.