On Decentralised Clustering in Self-monitoring Networks

Piraveenan Mahendra rajah[1], Mikhail Prokopenko[2], Peter Wang[2], Don Price[3] *

[1] Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5000, Australia
[2] CSIRO ICT Centre, Locked bag 17, North Ryde, NSW 1670,Australia
[3] CSIRO Industrial Physics, PO Box 218, Lindfield, NSW 2070, Australia

contact author: mikhail.prokopenko@csiro.au

ABSTRACT

A Decentralised Adaptive Clustering (DAC) algorithm for multi-agent networks is contrasted with a Fixed-order Centralised Adaptive Clustering algorithm (FCAC). The clustering is done on sensor readings detected within a self-monitoring impact sensing network. Simulation results show that DAC algorithm scales well with increasing network and data sizes and in some cases outperforms FCAC algorithm. While the common-sense intuition suggests that centralised algorithm is always superior, we support the simulation results with a simple counter-example.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Algorithms, Design, Experimentation

Keywords

clustering, sensor networks, scalability

1. INTRODUCTION

Currently, Structural Health Monitoring (SHM) — detection and evaluation of the extent and severity of structural damage — is carried out in a very limited way in specific regions, generally using a small number of sensors connected to a centralised data logger or computer. Ultimately, large numbers of sensors will be required in the exploitation of structures, operating in harsh working environments and responding to various forms of damage and failures. Another key requirement of an autonomous SHM system is robustness: its performance must degrade "gracefully" rather than catastrophically when damage occurs. Scalability and performance verification are also key requirements.

CSIRO-NASA Ageless Aerospace Vehicle (AAV) project developed and examined several essential concepts for self-monitoring sensing and communication networks [4, 1, 2]. These concepts are being developed, implemented and tested in a Concept Demonstrator (CD): a hardware multi-cellular sensing and communication network whose aim is to detect and react to impacts by projectiles that, for a vehicle in space, might be micro-meteoroids or space debris. The structure of the CD is a hexagonal prism. A modular aluminium frame is covered by 220 mm x 200 mm, 1 mm thick aluminium panels, forming the outer skin of the structure: each panel comprises 4 cells. Every AAV cell contains 4 passive piezoelectric polymer sensors bonded to the aluminium panel in order to detect the elastic waves generated by impacts; and 2 digital signal processors, one of which acquires data from the sensors, while the other runs the agent software and controls the communications with its neighboring cells. Importantly, a cell communicates only with four immediate neighbors. The CD does not employ centralised controllers or communication routers. A stand-alone Asynchronous Simulator capable of simulating the CD dealing with some environmental effects such as particle impacts of various energies has been developed and used in the reported experiments.

Autonomous SHM systems require an efficient discovery of main trends or unusual patterns in sensor-data, and in the absence of centralised controllers, rely on emergence of dynamic reconfigurable clusters, with some cells taking the roles of "local hierarchs". Most existing algorithms for clustering focus on how to form clusters, given a file or database containing the items. Decentralisation creates the additional complication that, even if a correct classification can be determined with the incomplete information, the location of items belonging to a class also needs to be discovered [3]. Moreover, new events may require dynamic reconfiguration of clusters on the basis of local sensor signals, and the cluster algorithm should be robust in the face of changes caused by cells failures and repairs.

2. ADAPTIVE CLUSTERING

Our main goal is an evaluation of a simple clustering technique in a dynamic and decentralised setting, exemplified by the AAV sensor and communication network, in terms of scalability and convergence. The algorithm input can be described as a series (a flux) of impact energies detected at different times and locations, while the output is a set of non-overlapping clusters, each with a dedicated cluster-head (a cell) and a cluster map of its followers (the AAV cells which detected the impacts) in terms of their sensor-data and relative coordinates. A cluster-head is dynamically selected among the set of nodes and acts as a local coordinator of transmissions within the cluster. The Decentralised Adaptive Clus-

*A part of this work was supported by a CSIRO ICT Centre scholarship to the first author. The authors are also grateful to Dr. Edward Generazio and Dr. William Prosser (NASA Langley Research Center), other members of the AAV project, and computing staff of the School of EEE (University of Adelaide).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
AAMAS'05, July 25-29, 2005, Utrecht, Netherlands.
Copyright 2005 ACM 1-59593-094-9/05/0007 ...$5.00.

tering (DAC) algorithm is described elsewhere [2]: it involves a number of inter-agent messages notifying agents about their sensory data, and changes in their relationships and actions. For example, an agent may send a recruit message to another agent, delegate the role of cluster-head to another agent, initiate a new cluster, etc. Most of these and similar decisions are based on the clustering heuristic [3], and a dynamic offset range [2]. The recruiting is done periodically, affecting all agents with impact-energy values within a particular offset ε of the value x of this agent. The clustering heuristic determines if a cluster should be split in two, and the location of this split. If there are two clusters, the offset of each new cluster-head is adjusted in such a way that the cluster-head of the "smaller" agents (relative to the value x) can now reach up to, but not including, the "smallest" agent in the cluster of "larger" agents. Similarly, the cluster-head of "larger" agents can now reach down to, but not including, the "largest" agent (the cluster-head) of the cluster of "smaller" agents. A centralised version was also developed — Fixed-order Centralised Adaptive Clustering (FCAC).

The quality of clustering (both DAC and FCAC) is measured by the weighted average cluster diameter, used by Zhang et al. [5]. The average pair-wise distance D for a cluster C with points $\{x_1, x_2, \ldots, x_m\}$ is given by

$$D = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} d(x_i, x_j)}{m(m-1)/2} \ ,$$

where $d(x_i, x_j)$ is the Euclidean distance between x_i and x_j. The weighted average cluster diameter for k clusters is given by:

$$\bar{D} = \sum_{i=1}^{k} m_i(m_i - 1)D_i \ / \ \sum_{i=1}^{k} m_i(m_i - 1) \ ,$$

where m_i is the number of elements in the cluster C_i with the pair-wise distance D_i. Neither decentralised nor centralised algorithm guarantees a convergence minimising the diameter $\bar{D}$. In fact, DAC may give different clusterings for the same set of agent values, depending on the physical locations of the impact points. The reason is a different communication flow affecting the adjustment of the offsets. The scope of agents involved in the clustering heuristic depends on the order of message passing, which in turn depends on the physical locations of impacts. The adjusted offsets determine which agents can be reached by a cluster-head, and this will affect the result of clustering. Therefore, for any set of agent values, there are certain sequences of events which yield better clustering results than others.

The developed centralised algorithm does not simulate all possible sequences of events, working with a *fixed-order* sequence. A random order, obviously, may not initiate the "best" ordering of events and yield the best clustering for a particular data set. In other words, centralisation of sensor-data is not a guarantee of a superior performance, and processing of all permutations is prohibitive even for a very small number of elements.

3. A COMPARISON AND CONCLUSIONS

Random numbers were used as agent values in simulations carried out to compare DAC and FCAC algorithms. The scalability and convergence analysis considered two scenarios. The first scenario (65 runs) kept the AAV grid array size constant, while increasing the number of impacts detected within it (i.e., the density of impacts was increasing). It was observed [2] that the relative performance of DAC in terms of $\bar{D}$ scales well and decreases "gracefully" with the number of impacts. DAC also needs relatively more and more executions of clustering heuristics to stabilise than the FCAC algorithm. This is expected, as the clustering heuristic needs to be invoked in many different agents with limited information.

The second scenario (50 runs), on the contrary, fixed the number of impacts, while increasing the grid size (i.e., the density of impacts was decreasing). In this case, not only the DAC algorithm scales very well with respect to the AAV array sizes, but it begins to outperform the FCAC algorithm when the array becomes larger. In addition, DAC begins to require fewer executions of clustering heuristics than FCAC: the larger data samples have more possible orderings and FCAC has a lesser chance to process the best one [2].

Here we provide such a sequence, as a counter-example to the common-sense intuition suggesting that centralised algorithm is always superior. The sequence is not the shortest possible (it includes 20 elements), but it pinpoints the reason very clearly. Consider the data set with agent values $\{0, 9, 21, 23, 38, 44, 49, 66, 87, 90, 96, 97, 102, 108, 120, 121, 128, 143, 151, 156\}$. The algorithms returned the following clustering results:

FCAC Clustering: $\{\{0\}, \{9\}, \{21, 23\}, \{38, 44, 49\}, \{66\}, \{87, 90, 96, 97, 102\}, \{118\}, \{120, 121, 128\}, \{143, 151, 156\}\}$ with the weighted average cluster diameter $\bar{D}_{FCAC} = 14.0$;

DAC Clustering: $\{\{0\}, \{9\}, \{21, 23\}, \{38, 44, 49\}, \{66\}, \{87, 90\}, \{96, 97, 102\}, \{118\}, \{120, 121, 128\}, \{143, 151, 156\}\}$ with the weighted average cluster diameter $\bar{D}_{DAC} = 11.571$.

In other words, this particular ordering of data gave worse results for the FCAC algorithm. Since it is impossible to simulate all $20!$ orderings for the FCAC algorithm, we simulated just 20 different orderings by randomly shuffling the data. Results included 8 outcomes with $\bar{D}_{FCAC} = 14.0$, 1 outcome with $\bar{D}_{FCAC} = 13.667$, and 11 outcomes with $\bar{D}_{FCAC} = 11.571$. Thus, the majority of FCAC runs converged on the better result shown by the DAC algorithm, confirming that data orderings make a difference to the performance of the FCAC algorithm. In short, FCAC algorithm is only an approximation of an ideal centralised clustering; the latter being infeasible in dynamic decentralised multi-agent networks.

The observation that DAC algorithm scales reasonably well with respect to array sizes and the number of impacts, and is robust in the face of a spatiotemporal impact flux [2], provides a very good support for deploying other, more sophisticated algorithms in dynamic impact sensing networks.

4. REFERENCES

[1] M. Foreman, M. Prokopenko, and P. Wang. Phase Transitions in Self-organising Sensor Networks. Proceedings of the 7th European Conference on Artificial Life, 781–791, LNAI 2801, Springer Verlag, 2003.

[2] P. Mahendra rajah, M. Prokopenko, P. Wang, and D. Price. Towards Adaptive Clustering in Self-monitoring Multi-Agent Networks. Proceedings by the 9th International Conference on Knowledge Based and Intelligent Information and Engineering Systems, Melbourne, 2005 (to appear).

[3] E. Ogston, B. Overeinder, M. Van Steen, and F. Brazier. A Method for Decentralized Clustering in Large Multi-Agent Systems. Proceedings of the 2nd International Joint Conference on Autonomous Agent and Multi Agent Systems, 798–796, 2003.

[4] D. Price, A. Scott, G. Edwards, A. Batten, A. Farmer, M. Hedley, M. Johnson, C. Lewis, G. Poulton, M. Prokopenko, P. Valencia, P. Wang. An Integrated Health Monitoring System for an Ageless Aerospace Vehicle. Proceedings of the 4th International Workshop on Structural Health Monitoring, Stanford, September 2003.

[5] T. Zhang, R. Ramakrishnan, M. Livny. BIRCH: A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery*, 1(2), 141-182, 1997.