On Convergence of Dynamic Cluster Formation in Multi-Agent Networks

Mikhail Prokopenko[1], Piraveenan Mahendra rajah[2], Peter Wang[1]

[1]CSIRO Information and Communication Technology Centre
Locked bag 17, North Ryde, NSW 1670, Australia
[2]2/177, Jeffcott Street, North Adelaide, SA 5006, Australia

Abstract. Efficient hierarchical architectures for reconfigurable and adaptive multi-agent networks require dynamic cluster formation among the set of nodes (agents). In the absence of centralised controllers, this process can be described as self-organisation of dynamic hierarchies, with multiple cluster-heads emerging as a result of inter-agent communications. Decentralised clustering algorithms deployed in multi-agent networks are hard to evaluate precisely for the reason of the diminished predictability brought about by self-organisation. In particular, it is hard to predict when the cluster formation will converge to a stable configuration. This paper proposes and experimentally evaluates a predictor for the convergence time of cluster formation, based on a regularity of the inter-agent communication space as the underlying parameter. The results indicate that the generalised "correlation entropy" K_2 (a lower bound of Kolmogorov-Sinai entropy) of the volume of the inter-agent communications can be correlated with the time of cluster formation, and can be used as its predictor.

1 Introduction

Dynamic creation and maintenance of "optimal" hierarchies in large dynamic networks is a well-recognised challenge. It appears in many different contexts, e.g., as dynamic hierarchies in Artificial Life [15], coalition formation in Agent-based Systems [17], decentralised clustering in Multi-Agent Systems [12], dynamic cluster formation in Mobile Ad Hoc Networks [10], etc. In this paper, we consider a sub-problem from this class: dynamic cluster formation in a sensor and communication network without centralised controllers. This process can be described as *self-organisation* of dynamic hierarchies, with multiple cluster-heads emerging as a result of inter-agent communications. Importantly, the emphasis is on rules of interactions (or communication protocols) between the engaged lower-level entities (cells, agents, network nodes, etc.) and the structures and patterns emerging at a higher-level (multi-cellular boundaries, multi-agent coalitions, local hierarchs or cluster-heads, etc.).

In general, the clustering of sensor-data aims at grouping entities with similar characteristics together so that main trends or unusual patterns may be discovered. Self-organising cluster formation in multi-agent networks/systems has two specific primary challenges: a) decentralised clustering: even if a correct classification can be determined with the incomplete information available, the location of items belonging to a class also needs to be discovered, "data is widely distributed, data sets are volatile, or data items cannot be compactly represented" [12], and; b) dynamic (on-line) clustering: new events may require reconfiguration of clusters: the resulting patterns or clusters have to be constantly refined. This requires efficient algorithms for decentralised sensor-data clustering in a distributed multi-agent system. A method for grouping networked agents with similar objectives or data without collecting them into a centralised database is presented by Ogston et al. [12], and shows very good scalability and speed in comparison with the k-means clustering algorithm. It employs a heuristic for breaking large clusters when required, and a sophisticated technique dynamically matching agents objectives, represented as connections in the multi-agent network.

However, decentralised clustering algorithms deployed in multi-agent networks are hard to evaluate precisely for the reason of the diminished predictability brought about by self-organisation. In particular, it is hard to predict when the cluster formation will converge to a stable configuration. Such a predictive ability is, however, important for deciding whether clusters will form in time for multi-agent diagnostics, being a prerequisite for the overall damage propagation prognosis. The specific objective of this paper is an identification and evaluation of potential predictors for the convergence time of dynamic cluster formation.

In achieving this goal, we analyse two levels of multi-agent dynamics: macro-level, where coordination patterns form and can be observed, and micro-level, where the inter-agent messages are exchanged, creating a multi-agent communication space. We consider irregularity of the inter-agent communication space, and propose it as a possible predictor for our task. This predictor is estimated via the generalised "correlation entropy" K_2 of the underlying time series: the traffic volume of inter-agent communications. The estimates are shown to be correlated with the convergence time of cluster formation.

The experiments required to evaluate the predictor were carried out on a self-monitoring sensor and communication network developed CSIRO-NASA "Ageless" Aerospace Vehicle (AAV) project, in the context of Structural Health Management (SHM). The AAV project is briefly described in the next section, followed by a simplified version of a decentralised adaptive clustering algorithm developed for evaluation purposes. Section 3 presents the proposed predictor for the convergence time of cluster formation, followed by a discussion of the obtained results and future work.

2 Adaptive Clustering in Self-organising SHM Networks

Structural health monitoring and management of complex, safety-critical structures such as aerospace vehicles will ultimately require the development of intelligent networks systems that can process the data from large numbers of sensors; evaluate and diagnose detected damage; form a prognosis for the damaged structure; make decisions regarding response to or repair of the damage; initiate the required actions and monitor their effectiveness [13, 3]. Recently, several essential concepts for self-organising SHM networks as well as their desirable characteristics, such as robustness, reliability and scalability, have been identified in the literature [13, 14]. Some of these concepts are being developed, implemented and tested in the AAV Concept Demonstrator (AAV-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. A stand-alone Asynchronous Simulator capable of simulating the AAV-CD dealing with some environmental effects such as particle impacts of various energies has been developed and used in the reported experiments. The damage sensing network may consists of "cells" (agents) that not only form a physical shell ("skin") for a structure (e.g., an aerospace vehicle), but also have passive sensors detecting elastic waves generated in the "skin" by impacts; and electronic modules, acquiring data from the sensors, running the agent software and controlling the communications with its neighbouring cells. Importantly, a cell should communicate only with immediate neighbours, eliminating single critical points of failure: all data are processed locally, and only information relevant to other regions of the structure is communicated.

Single cells may detect impacts and triangulate their locations, while collections of cells may solve more complex tasks. Some responses could be purely local, while some may require emergence of dynamic reconfigurable structures, with some cells taking the roles of "local hierarchs". A cluster-head may be dynamically selected among the set of nodes and become a local coordinator of transmissions within the cluster. A typical SHM task may require impact-data clusters, logically grouping the cells which detected impacts with energies within a certain band (e.g., non-critical impacts). Moreover, clusters would form and re-form when new damage is detected on the basis of local sensor signals. Importantly, a cluster formation algorithm should be robust to changes caused by new impacts, cells' failures and possible repairs.

As pointed out earlier, our main goal is not a new clustering method *per se*, but rather an analysis of a representative clustering technique in a dynamic and decentralised multi-agent setting, exemplified by the AAV sensor and communication network, *in terms of predictability of its convergence time*. There are some important specific details of our experimental setup which may be relevant to other multi-agent networks: a particular communication infrastructure where each cell is connected only to immediate neighbours; constraints on the communication bandwidth; dynamic scenarios where density of events may vary in time and space; a decentralised architecture without absolute coordinates or id's of individual cells on a large-scale multi-cellular skin. To stay within a generic framework, we abstracted away almost all sensor-data features. For example, instead of considering time-domain or frequency-domain impact data, detected and/or processed by cell sensors [13], we represent a cell sensory reading with a single aggregated value ("impact-energy"), define "differences" between cells in terms of this value, and attempt to cluster cells while minimising these "differences". This approach can be relatively easily extended to cases where "differences" are defined in a multi-dimensional space. In short, our focus is on inter-agent communications required by a decentralised clustering algorithm, dynamically adapting to changes, and the convergence time.

2.1 Dynamic Cluster Formation Algorithm

The algorithm input can be described as a series (a flux) of events (impacts) detected at different times and locations, while the output is a set of non-overlapping clusters, each with a dedicated cluster-head (a network cell) and a cluster map of its followers (cells which detected the impacts) in terms of their sensor-data and relative coordinates. The algorithm is described elsewhere [11] and 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, or declare "independence" by initiating a new cluster. Most of these and similar decisions are based on the clustering heuristic described by Ogston et al. [12], and a dynamic offset range. This heuristic determines if a cluster should be split in two, and the location of this split.

Each cluster-head (initially, each agent) broadcasts its *recruit* message periodically, with a broadcasting-period, affecting all agents with values within a particular dynamic offset ε of the impact-energy data x detected by this agent. Every *recruit* message contains the sensor-data of all current followers of the cluster-head with their relative coordinates (a cluster map). Under certain conditions, an agent, which is not a follower in any cluster, receiving a *recruit* message becomes a follower, stops broadcasting its own *recruit* messages and sends its information to its new cluster-head indicating its relative coordinates and the sensor-data x. However, there are situations when the receiving agent is already a follower in some cluster and cannot accept a recruit message by itself — a recruit disagreement. In this case, this agent *forwards* the received recruiting request to its present cluster-head. Every cluster-head waits for a certain period, collecting all such *forward* messages, at the end of which the clustering heuristic is invoked on the union set of present followers and all agents who *forwarded* their new requests.

Firstly, all n agents in the combined list are sorted in decreasing order according to their impact-energy value x. Then, a series of all possible divisions in the ordered set of agents is generated. That is, the first ordering is a cluster with all agents in it; the second ordering has the agent with the largest value in the first cluster and all other agents in the second cluster; and so forth (the n-th division has only the last n-th agent in the second cluster). For each of these divisions, the quality of clustering is measured by the total square error:

$$E_j^2 = \sum_{i=1}^{z} \sum_{x \in A_{i,j}} \|x - m_{i,j}\|^2 \ ,$$

where z is a number of considered clusters ($z = 2$ when only one split is considered), $A_{i,j}$ are the clusters resulting from a particular division and $m_{i,j}$ is the mean value of the cluster $A_{i,j}$. We divide E^2 values by their maximum to get a series of normalised values. Then we approximate the second derivative of the normalised errors per division:

$$f''(E_j^2) = (E_{j+1}^2 + E_{j-1}^2 - 2E_j^2) \ / \ h^2 \ ,$$

where $h = \frac{1}{n}$. If the peak of the second derivative is greater than some threshold for a division j, we split the set accordingly; otherwise, the set will remain as one cluster.

The cluster-head which invoked the heuristic notifies new cluster-heads about their appointment, and sends their cluster maps to them: a *cluster-information* message. When the clustering heuristic is applied, it may produce either one or two clusters as a result. If there are two clusters, the offset of each new cluster-head is modified. It is adjusted in such a way that the cluster-head of the “smaller” agents (henceforth, references like “larger” or “smaller” are 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. These adjusted offsets are sent to the new cluster-heads along with their cluster maps.

There are other auxiliary messages involved in the algorithm but importantly, the cluster formation is driven by three types: *recruit*, *cluster-information*, and *forward* messages. The first two types are periodic, while the latter type depends only on the degree of disagreements among cluster-heads. On the one hand, if there are no disagreements in clustering (for instance, if a clustering heuristic resulted in optimal splits even with incomplete data), then there is no need in *forward* messages. On the other hand, when cluster-heads frequently disagree on formed clusters, the *forward* messages are common. In short, it is precisely the number of *forward* messages traced in time — the traffic volume of inter-agent communications — that we hope may provide an underlying time series $\{v_t\}$ for our prognostic analysis, as it exhibits both periodic and chaotic features.

The quality of clustering is measured by the weighted average cluster diameter [23], but the algorithm does not guarantee a convergence minimising this criterion. In fact, it 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. Each time the clustering heuristic is executed in an agent, its offsets are either left alone or reduced. 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.

We conducted extensive preliminary simulations to determine whether the algorithm is robust and scales well in terms of the quality of clustering and convergence, measured by the number of times the clustering heuristic was invoked before stability is achieved with each data set [11]. Several scenarios were considered. The first scenario kept the network size constant, while increasing the number of impacts detected within it. The second scenario, on the contrary, fixed the number of impacts, while increasing the network size. In other words, the density of impacts was increasing in the first case, and decreasing in the second. Finally, we developed scenarios where impacts appear periodically, with varying periods. While the simulation results show that the algorithm converges and scales well in all cases, and in addition, is robust to dynamics of the sensor-data flux, the convergence time varies significantly (Figure 1), without obvious indicative patterns.

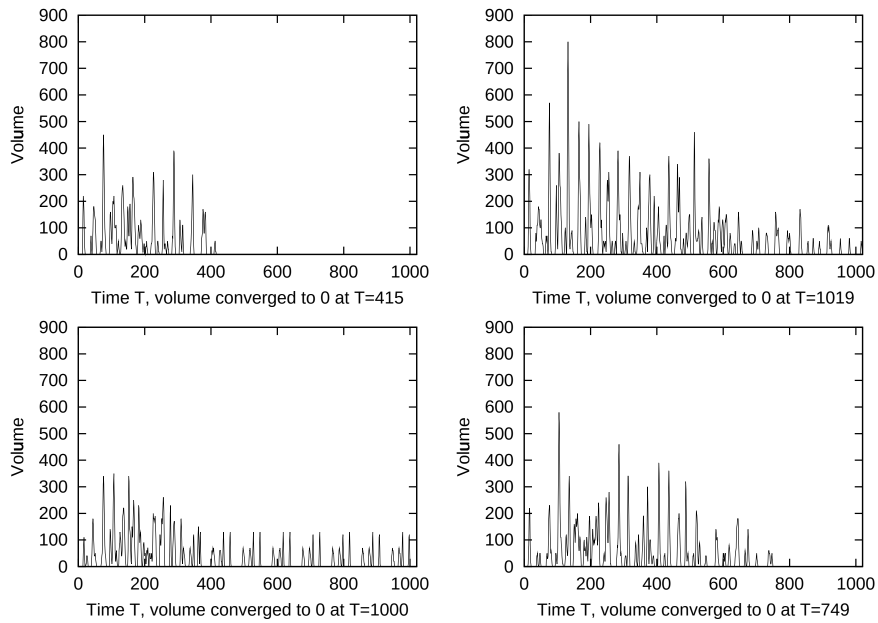

Fig. 1. Varying convergence times T_s for 4 different experiments, $1 \leq s \leq 4$.

In the remainder of the paper we focus on our main objective: prediction of the convergence time T, based on regularity of an initial segment $0, ..., \mathcal{D}$, where $\mathcal{D} < T$, of the "communication-volume" series $\{v(t)\}$, where $v(t)$ is the number of *forward* messages at time t.

3 The $K_q(\mathcal{D})$ predictor: Entropy of multi-agent communication-volume

The observed variability of different communication-volume time series may indicate that the underlying dynamics in the phase-space includes both unstable periodic and chaotic orbits, and an unstable fixed-point. It is known that in many experiments, time series often exhibit irregular behavior during an initial interval before finally settling into an asymptotic state which is non-chaotic [1] — in our case, eventually converging to a fixed-point ($v_T = 0$). The irregular initial part of the series may, nevertheless, contain valuable information: this is particularly true when the underlying dynamics is deterministic and exhibits *transient chaos* [1, 7]. We believe that the described algorithm for dynamic cluster formation, employing the clustering heuristic and adjustments of the offset ε, creates *multi-agent transient chaotic dynamics*.

Our plan is simple: for each experiment s, a) select an initial segment of length $\mathcal{D}$ of the time series, $\{v_s^{\mathcal{D}}\}$; b) estimate its generalised entropy $K_q(\mathcal{D})_s$ for a range of estimation-dependent parameters (see the description below). Then, c) given the estimates $K_q(\mathcal{D})_s$ for all the experiments, correlate them with the observed convergence times T_s, e.g., by using a linear regression $T = a + bK_q(\mathcal{D})$ and the correlation coefficient ρ between the series $\{T_s\}$ and $\{K_2(\mathcal{D})_s\}$. This would allow us to predict the time T_s of convergence to $v_s(T_s) = 0$, as $T_s = a + bK_q(\mathcal{D})_s$.

A simple characterisation of the "regularity" of the communication space is provided by the auto-correlation function of an integer delay τ:

$$\gamma_s(\tau) \ = \ \sum_{t=\tau+1}^{\mathcal{D}} \left[v_s(t-\tau) - \overline{v_s}\right] \left[v_s(t) - \overline{v_s}\right] / \sum_{t=1}^{\mathcal{D}} [v_s(t) - \overline{v_s}]^2 \ . \tag{1}$$

The auto-correlation is obviously limited to measuring only linear dependencies, and we consider instead a more general and elaborate approach. One classical measure is the Kolmogorov-Sinai (KS) entropy, also known as metric entropy [8, 9, 18]: it is a measure for the rate at which information about the state of the system is lost in the course of time. In other words, it is an entropy per unit time, or an "entropy rate". Suppose that the $d-$dimensional phase space is partitioned into boxes of size r^d. Let $P_{i_0 \ldots i_{d-1}}$ be the joint probability that a trajectory is in box i_0 at time 0, in box i_1 at time Δt, ..., and in box i_{d-1} at time $(d-1)\Delta t$, where Δt is the time interval between measurements on the state of the system (in our case, we may assume $\Delta t = 1$, and omit the limit $\Delta t \to 0$ in the following definitions). The KS entropy is defined by

$$K = - \lim_{\Delta t \to 0} \lim_{r \to 0} \lim_{d \to \infty} \frac{1}{d \Delta t} \sum_{i_0 \ldots i_{d-1}} P_{i_0 \ldots i_{d-1}} \ln P_{i_0 \ldots i_{d-1}} \ , \tag{2}$$

and more precisely, as a supremum of K on all possible partitions. This definition has been generalised to the order-q Rényi entropies K_q [16]:

$$K_q = - \lim_{\Delta t \to 0} \lim_{r \to 0} \lim_{d \to \infty} \frac{1}{d \Delta t (1-q)} \ln \sum_{i_0 \ldots i_{d-1}} P^q_{i_0 \ldots i_{d-1}} \ . \tag{3}$$

It is well-known that $K = 0$ in an ordered system, K is infinite in a random system, and K is a positive constant in a deterministic chaotic system. Grassberger and Procaccia [4] considered the "correlation entropy" K_2 in particular, and capitalised on the fact $K \geq K_2$ in establishing a sufficient condition for chaos $K_2 > 0$. The Grassberger and Procaccia (GP) algorithm estimates the entropy K_2 as follows:

$$K_2 = \lim_{r \to 0} \lim_{d \to \infty} \lim_{N \to \infty} \ln \frac{C_d(N,r)}{C_{d+1}(N,r)} \ , \tag{4}$$

where $C_d(r)$ is the correlation integral:

$$C_d(N,r) = \frac{1}{(N-1)N} \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta(r - \|\boldsymbol{V_i} - \boldsymbol{V_j}\|) \ . \tag{5}$$

Here Θ is the Heaviside function (equal to 0 for negative argument and 1 otherwise), and the vectors $\boldsymbol{V_i}$ and $\boldsymbol{V_j}$ contain elements of the observed time series $\{v(t)\}$, "converting" or "reconstructing" the dynamical information in one-dimensional data to spatial information in the d-dimensional embedding space: $\boldsymbol{V_k} = (v_k, v_{k+1}, v_{k+2}, \ldots, v_{k+d-1})$ [19]. The norm $\|\boldsymbol{V_i} - \boldsymbol{V_j}\|$

is the distance between the vectors in the d-dimensional space, e.g., the maximum norm [20]:

$$\|\boldsymbol{V_i} - \boldsymbol{V_j}\| = \max_{\tau=0}^{d-1}(v_{i+\tau} - v_{j+\tau}) \tag{6}$$

Put simply, $C_d(r)$ computes the fraction of pairs of vectors in the d-dimensional embedding space that are separated by a distance less than or equal to r. In order to eliminate auto-correlation effects, the vectors in Equation (5) should be chosen to satisfy $|i - j| > L$, for some positive L, and at the very least $i \neq j$ [21]. Since we consider only an initial segment of the times series, we simply set $N = \mathcal{D}$ in the Equation (5), estimating the entropy as

$$K_2(d, r, \mathcal{D}) = \ln \frac{C_d(\mathcal{D}, r)}{C_{d+1}(\mathcal{D}, r)} \ . \tag{7}$$

Now we only need to identify the embedding dimension $\hat{d}$ and the distance $\hat{r}$ which maximise the correlation coefficient for s experiments, $\rho(\{T_s\}, \{K_2(d, r, \mathcal{D})_s\})$, over a range of d and r, and designate

$$K_2(\mathcal{D})_s = K_2(\hat{d}, \hat{r}, \mathcal{D})_s. \tag{8}$$

At this stage we need to make a comment on the correlation dimension. Within certain ranges of r and d, the correlation integral $C_d(r)$ may be proportional to some power of r, $C_d(r) \sim r^\nu$ [5]. This power ν is called the correlation dimension. If the dynamical process is unfolded by choosing a sufficiently large $d > d_\nu$, a typical slope of the plot $\ln C_d(r)$ versus $\ln r$ becomes independent of d. We observed (see the next section) that this minimum embedding dimension d_ν is not necessarily the embedding dimension $\hat{d}$ maximising the predictor $K_2(\mathcal{D})$, but a possible connection is intriguing.

The correlation dimension provides useful information about the local structure of the process and is an effective measure of its (possibly fractal) size: in particular, a random process has an "infinite" correlation dimension (its orbit is not expected to have any spatial structure). In contrast, the correlation dimension for a periodic orbit is 1, while it could be higher for some non-regular processes. A non-integer $\nu < 1$ is an indication of a strange chaotic attractor [5]. It is worth pointing out that the GP algorithm can be used to estimate the correlation dimension of underlying chaotic transients [1].

4 Experimental Results

The experiments included $s = 1, ..., 20$ runs of the clustering algorithm, tracing the communication-volume time series $\{v_t\}$. As expected, the auto-correlation function $\gamma_s(\tau)$, Equation (1), did not advance us in our experiments: the highest correlation coefficient between convergence times T_s and auto-correlations $\gamma_s(\tau)$, for a range of delays τ, was only 0.52.

We then selected an initial segment $\mathcal{D} = 400$, and computed correlation integrals $C_d(400, r)$ for a wide range of embedding dimensions ($d < 98$) and distances ($1 \leq r \leq 1000$, a median standard deviation of $\{v\}_s$ being about 100). A plot $\ln C_d(r)$ versus $\ln r$ is shown in Figure 2, illustrating the time series depicted in the top-left of Figure 1 (a quickly converged series, $T = 415$). We can observe three well-known regions: 1) the lower region distorted by fluctuations due to the small number of points, 2) a linear "scaling" region where the power law $C_d(r) \sim r^\nu$ holds, and 3) the upper region distorted due to the finite size of the process. There are also anomalous shoulders in the correlation integral due to remaining autocorrelation in the time-series data [22]. We observed that quickly converged series have an earlier onset of the upper region than slowly converged series. The plot strongly indicates a transient multi-mode process, and suggests the possibility of extracting meaningful predictors $K_2(\mathcal{D})_s$, specified by Equations (7)-(8).

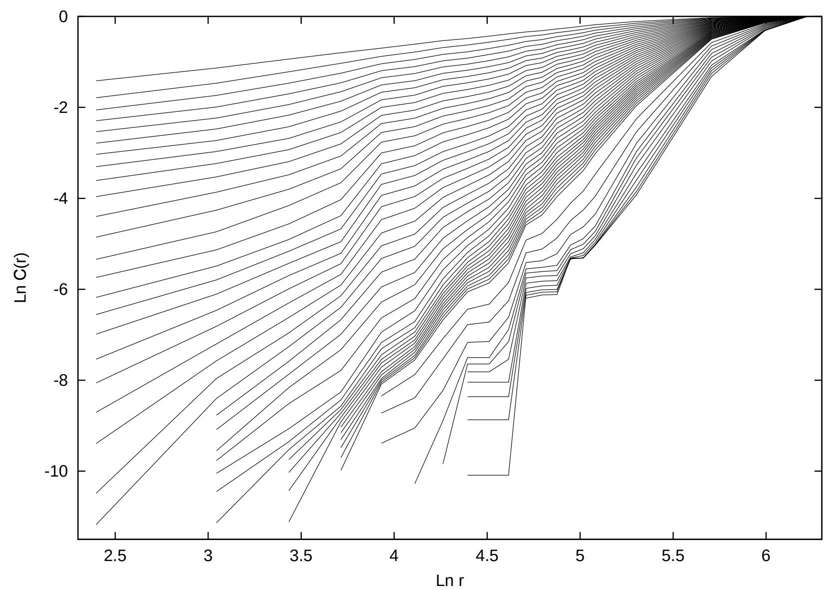

Fig. 2. A plot $\ln C_d(r)$ versus $\ln r$ for a quickly converged series, $T = 415$. Embedding dimensions are shown for every d between 1 (from top-left corner) and 40, and for every 5th d between 45 and 95 (towards bottom-right corner).

Given data of s experiments: the 3-dimensional array $K_2(d, r, \mathcal{D})_s$ for varying d and r, and each s, the correlation coefficient $\rho(\{T_s\}, \{K_2(d, r, \mathcal{D})_s\})$ was determined for the range of d and r. It is shown in Figure 3, clearly reaching maximum at embedding dimensions $6 \leq d \leq 10$, almost uniformly for all the distances r. The maximum ($\rho = 0.898345$) was attained at $\hat{d} = 6$ and $\hat{r} = 41$, and is a very encouraging correlation value. Figure 4 shows the linear regression between $\{T_s\}$ and $\{K_2(\mathcal{D})_s\}$, where the latter is selected according to the Equation (8) for the identified $\hat{d}$ and $\hat{r}$.

5 Conclusions and Future Work

We considered decentralised and dynamic cluster formation in multi-agent sensor and communication networks, proposed and experimentally evaluated a predictor for the convergence time of cluster formation. The predictor $K_2(\mathcal{D})$ is based on the generalised "correlation entropy" (a lower bound of Kolmogorov-Sinai entropy) of the volume of the inter-agent communications. The results indicate that $K_2(\mathcal{D})$ can be well correlated with the time of cluster formation. The predictor $K_q(\mathcal{D})$ can also be considered for other orders q, and this work is ongoing.

The dynamic cluster formation may be interpreted in self-referential terms: inter-agent messages contribute to emergence of cluster hierarchies at macro-level, and at the same time are significantly influenced by them at micro-level. Such an interdependency can also be characterised in terms of tangled hierarchies exhibiting Strange Loops [6]: "an interaction between levels in which the top level reaches back down towards the bottom level and influences it,

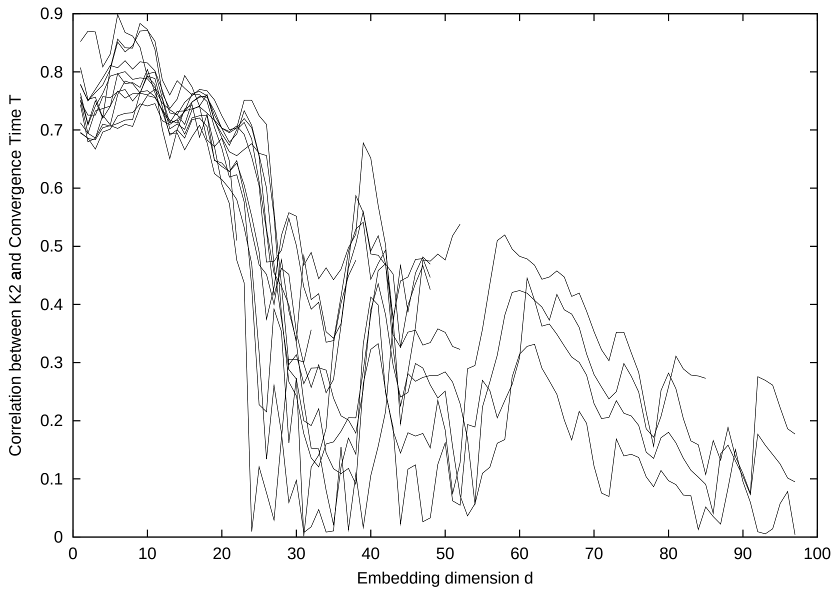

Fig. 3. The correlation coefficient ρ between the series $\{T_s\}$ and $\{K_2(\mathcal{D})_s\}$, for a range of distances r. The maximum ($\rho = 0.898345$) is at $\hat{d} = 6$ and $\hat{r} = 41$.

while at the same time being itself determined by the bottom level". The observed *multi-agent transient chaotic dynamics* may appear precisely due to this self-referentiality.

The performance of the predictor $K_2(\mathcal{D})$ provides a very good support for deploying other, more sophisticated algorithms in the sensing networks. The density-based algorithms may particularly be relevant in our application: e.g., DBSCAN algorithm would allow us to discover clusters with arbitrary shape [2].

Acknowledgements The authors are grateful to Ed Generazio and William Prosser (NASA Langley Research Center), Daniel Polani (University of Hertfordshire), Frank Horowitz (CSIRO Exploration and Mining), and to other members of the AAV team, especially Don Price and Tony Farmer (CSIRO Industrial Physics), for their encouragement and insightful comments.

References

1. Dhamala, M., Lai, Y.C., Kostelich, E.J. Analyses of transient chaotic time series. *Physical Review E*, 64, 056207, 1–9, 2001.
2. Ester, M., Kriegel, H., Sander, J., and Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. The 2nd International Conference on Knowledge Discovery and Data Mining, 226-231, 1996.
3. Foreman, M., Prokopenko, M., Wang, P. Phase Transitions in Self-organising Sensor Networks. Banzhaf, W., Christaller, T., Dittrich, P., Kim, J.T. Ziegler, J. (Eds.) *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life*, 781–791, LNAI 2801, Springer, 2003.
4. Grassberger, P. and Procaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Physical Review A*, 28(4):2591, 1983.
5. Grassberger, P. and Procaccia, I. Characterization of strange attractors. *Physical Review Letters*, 50:346–349, 1983.
6. Hofstadter, D.R. *Godel, Escher, Bach: An Eternal Golden Braid*. New York: Vintage Books, 1989.
7. Jánosi, I.M., and Tél, T. Time series analysis of transient chaos. *Physical Review E*, 49(4):2756–2763, 1994.

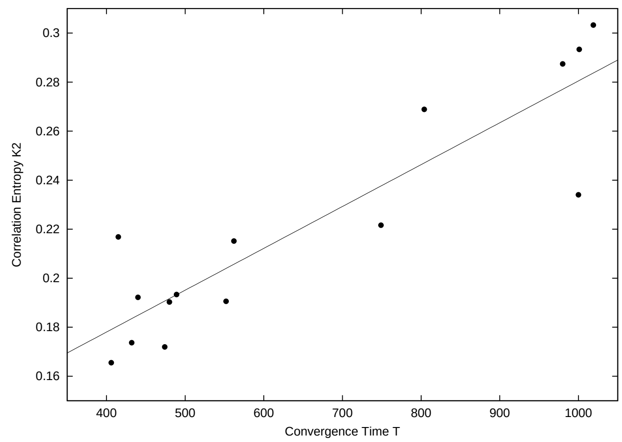

Fig. 4. The linear regression between $\{T_s\}$ and $\{K_2(\mathcal{D})_s\}$.

8. Kolmogorov, A.N. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Doklady Akademii Nauk SSSR*, 119:861-864 (Russian), 1958.
9. Kolmogorov, A.N. Entropy per unit time as a metric invariant of automorphisms. *Doklady Akademii Nauk SSSR*, 124:754–755 (Russian), 1959.
10. Lin R., and Gerla, M. Adaptive Clustering for Mobile Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 1265–1275, September 1997.
11. Mahendra rajah, P., Prokopenko, M., Wang, P., Price, D.C. Towards Adaptive Clustering in Self-monitoring Multi-Agent Networks. The 9th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (KES-2005), Melbourne, Australia, September 2005.
12. Ogston E., Overeinder, B., Van Steen, M., and Brazier, F. A Method for Decentralized Clustering in Large Multi-Agent Systems. The 2nd International Joint Conference on Autonomous Agent and Multi Agent Systems, 798–796, 2003.
13. Price, D.C., Scott, D.A., Edwards, G., Batten, A., Farmer, A.J., Hedley, M., Johnson, M., Lewis, C., Poulton, G., Prokopenko, M., Valencia, P., Wang, P. An Integrated Health Monitoring System for an Ageless Aerospace Vehicle. 4th Int'l Workshop on Structural Health Monitoring, Stanford, 2003.
14. Prokopenko, M., Wang, P.., Price, D.C., Valencia, P., Foreman, M., Farmer, A.J. Self-organising Hierarchies in Sensor and Communication Networks. To appear in *Artificial Life*, Special issue on Dynamic Hierarchies, 2005.
15. Rasmussen, S., Baas, N.A., Mayer, B., Nilsson, M., and Olesen, M.W. Ansatz for Dynamical Hierarchies. *Artificial Life*, vol. 7, 4, 2001.
16. Rényi, A. *Probability theory*. North-Holland, Amsterdam, 1970.
17. Sandholm, T. and Lesser, V. Coalition Formation among Bounded Rational Agents. The 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montreal, Canada, 662-669.
18. Sinai, Ya.G. On the concept of entropy of a dynamical system. *Doklady Akademii Nauk SSSR*, 124:768-771 (Russian), 1959.
19. Takens, F. Detecting strange attractors in turbulence. *Dynamical systems and turbulence*, LNM Vol.898, Springer, Berlin, 1981.
20. Takens, F. Invariants related to dimension and entropy. *Atas do 13* Colóquio Brasiliero do Matemática, Rio de Janeiro, 1983.
21. Theiler, J. Spurious dimension from correlation algorithms applied to limited time-series data. *Physical Review A*, 34(3):2427–2432, 1986
22. Theiler, J. Estimating fractal dimension. *J. of the Optical Society of America A*, 7(6):1055, 1990.
23. Zhang, T., Ramakrishnan, R., Livny, M. BIRCH:A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery*, 1(2), 141-182, 1997.