Communication Diversity in Particle Swarm Optimizers

Marcos Oliveira[1(✉)], Diego Pinheiro[1(✉)], Bruno Andrade[2],
Carmelo Bastos-Filho[3(✉)], and Ronaldo Menezes[1(✉)]

[1] BioComplex Laboratory, Florida Institute of Technology, Melbourne, USA
{moliveira,dsilva,rmenezes}@biocomplexlab.org
[2] Universidade Federal de Goiàs, Goiàs, Brazil
brunoandrade@inf.ufg.br
[3] Universidade de Pernambuco, Recife, Brazil
carmelo.filho@upe.br

Abstract. Since they were introduced, Particle Swarm Optimizers have suffered from early stagnation due to premature convergence. Assessing swarm spatial diversity might help to mitigate early stagnation but swarm spatial diversity itself emerges from the main property that essentially drives swarm optimizers towards convergence and distinctively distinguishes PSO from other optimization techniques: *the social interaction between the particles*. The swarm *influence graph* captures the structure of particle interactions by monitoring the information exchanges during the search process; such graph has been shown to provide a rich overall structure of the swarm information flow. In this paper, we define swarm *communication diversity* based on the component analysis of the swarm influence graph. We show how communication diversity relates to other measures of swarm spatial diversity as well as how each swarm topology leads to different communication signatures. Moreover, we argue that swarm *communication diversity* might potentially be a better way to understand early stagnation since it takes into account the (social) interactions between the particles instead of properties associated with individual particles.

Keywords: PSO · Swarm assessment · Premature convergence · Early stagnation analysis · Component graph analysis

1 Introduction

Particle Swarm Optimization (PSO) is a computational intelligence technique inspired by the social behavior of bird flocks and used to solve optimization problems [4]. The technique consists of a population (swarm) of simple reactive agents (particles) interacting among themselves while exploring the search space by seeking the best solutions. The communication between the particles plays an important role on the swarm behavior [8]. Such aspect may be seen, for

instance, when the swarm topology is analyzed. The topology defines the neighborhood of each particle thus impacting how the information is shared within the swarm. Less connected topologies slow down the information flow since information is transmitted indirectly through intermediary particles [5]. Conversely, highly-connected topologies decrease the number of these intermediaries creating a tendency for the swarm to move quickly towards local optima. These different behaviors are also intimately related to the exploration-exploitation balance in the swarm [6].

Although communication among particles is an essential aspect of swarm behavior, many researchers only focus on the final result of the interactions between them such as using the particles properties (*e.g.* their positions, velocities, etc.) to assess the swarm behavior [13,16]. These approaches bring interesting findings but they are actually analyzing the final results of the particles interactions and not the *information flow* of the swarm execution. Some of the few works addressing information flow introduced the concept of swarm influence graph that captures the actual flow of information between the particles [9,10].

The analyses performed with these approaches allow a better understanding of the swarm search mode. However, they do not capture the *dynamics* in the information flow during the algorithm execution, but only analyze a snapshot from a given iteration. The analysis of swarm dynamics can promote the comprehension of its behavior which leads, for example, to swarm stagnation.

In this work we define swarm *communication diversity* using component analysis of the influence graph. We show how our definition relates to swarm behavior and how each swarm topology leads to different communication signatures. We argue that swarm communication diversity is a more suitable analysis of the swarm because the social interactions of the particles are included, instead of only properties associated with individual particles.

2   PSO, Diversity and Particle Interactions

Particle Swarm Optimization is a stochastic, population-based optimization technique inspired by the social behavior of bird flocks and is composed by particles that move around the search space [4]. The equations that control the swarm are not included in this paper but an interested reader should refer to the standard PSO paper [1]. The particles can only share information with the ones in their neighborhood defined by the swarm topology consisting of a graph in which two nodes (particles) are linked if they are allowed to share information to each other. The topology influences the flow of information between particles. For instance, when the average distance between nodes is too short, the swarm tends to move quickly towards the best solution found in earlier iterations which usually leads to a fast convergence to the global optimum in unimodal problems, but with the caveat of the possibility to prematurely reach a local optimum, specially in multimodal problems [1,5]. In this sense, communication topologies with fewer connections may yield better results. Since the information spreads slowly, the swarm has a better chance of exploring different regions of the search space.

In a global topology, the swarm shares the same (social) memory because particles are directly connected and any particle can be the information spreader to all the other particles. Conversely, the particles have different neighbors in local topologies and hence their social memories are different [1]. For instance, in a ring topology, the particles can only communicate with two other particles and although it prevents premature attraction of the whole swarm to a single location due to slow spread of information, it brings the inconvenience of a slow convergence time [1]. These extreme behaviors motivate considering topologies that balance their strengths and weaknesses such as the von Neumann topology [1,5].

2.1 Early Stagnation and Swarm Diversity

Convergence and stagnation are closely related terms often referred interchangeably although they subtly differ from each other. Convergence towards global optima tightly links to objective function improvement and is an appealing feature for any optimizer. Particularly in the case of PSO, premature convergence means that the swarm attained an equilibrium state which is often a local optima. Clearly, this is not desired since we would like the swarm to converge while effectively working in order to find better candidate solutions. In fact, many mechanisms to prevent premature convergence have been proposed [2,14,17].

Stagnation is often the result of an exploration-exploitation imbalance that causes the search space not to be explored adequately. This imbalance led researchers to believe early stagnation strongly relates to swarm spatial diversity and that consequently the assessment of such diversity is a way prevent undesired stagnation [14]. In this sense, PSO would ideally start the optimization process in exploration mode with a high diversity and, as the search space is adequately explored, it would initiate the exploitation mode focusing its efforts on smaller areas of the search space with a lower diversity.

Metrics for assessing swarm diversity are mostly spatial in nature and quantify swarm diversity as the degree of dispersion of particles around a given swarm center [12]. We enumerate the main diversity metrics as hereafter described. Let $|S|$ be the size of swarm S, $\boldsymbol{x}_i$ the position of particle i, $f[\cdot]$ the fitness function, $\boldsymbol{x}_{best}(t)$ the position of the best particle at iteration t, $\bar{f}(t)$ the average fitness at iteration t, and $d(\boldsymbol{p}, \boldsymbol{q})$ the euclidean distance between $\boldsymbol{p}$ and $\boldsymbol{q}$.

1. Aggregation degree:
$$AD(t) = \frac{f(\boldsymbol{x}_{best}(t))}{\bar{f}(t)}. \tag{1}$$

2. Normalized average distance around the swarm center using the swarm diameter or radius:
$$\mathcal{D}_{\bar{\boldsymbol{x}}}^{D}(t) = \frac{\mathcal{D}_{\bar{\boldsymbol{x}}}(t)}{D(t)}, \quad \mathcal{D}_{\bar{\boldsymbol{x}}}^{R}(t) = \frac{\mathcal{D}_{\bar{\boldsymbol{x}}}(t)}{R(t)}. \tag{2}$$

3. Average distance around the swarm center $\bar{x}$ [7]:

$$\mathcal{D}_{\bar{x}}(t) = \frac{1}{|S|}\sum_{i=1}^{|S|}d\big(x_i(t), \bar{x}(t)\big). \tag{3}$$

4. Average of the average distance around all particles in the swarm:

$$\mathcal{D}_{all}(t) = \frac{1}{|S|}\sum_{i=1}^{|S|}\mathcal{D}_{x_i}(t). \tag{4}$$

5. Diameter, the maximum distance between the positions of any two particles i and j in the swarm:

$$D(t) = \max_{i \neq j \in [1, |S|]} d\big(x_i(t), x_j(t)\big). \tag{5}$$

6. Radius, the maximum distance between the swarm center $\bar{x}$ and any particle i in the swarm:

$$R(t) = \max_{i \in [1, |S|]} d\big(x_i(t), \bar{x}(t)\big). \tag{6}$$

These metrics attempt to quantify swarm diversity as the amount of swarm expansion and dispersion. Both diameter, D, and radius, R, are highly sensitive to outliers since one single particle can greatly influence the result. Since they are employed as normalization factors, the same applies for the normalized average distance around the swarm center ($\mathcal{D}_{\bar{x}}^{D}$ or $\mathcal{D}_{\bar{x}}^{R}$). The average distance around the swarm center, $\mathcal{D}_{\bar{x}}$, as well as the average of the average distance around all particles in the swarm, $\mathcal{D}_{all}$, are less sensitive to outliers and while $\mathcal{D}_{\bar{x}}$ considers a swarm center $\bar{x}$, $\mathcal{D}_{all}$ takes the average of each particle as the center. The aggregation degree, AD, measures how close the best fitness $f(x_{best}(t))$ found at iteration t is from the average swarm fitness $\bar{f}(t)$ at iteration t. This supposedly approximates the degree of divergence among the new found candidate solutions thus measuring swarm diversity under the fitness perspective.
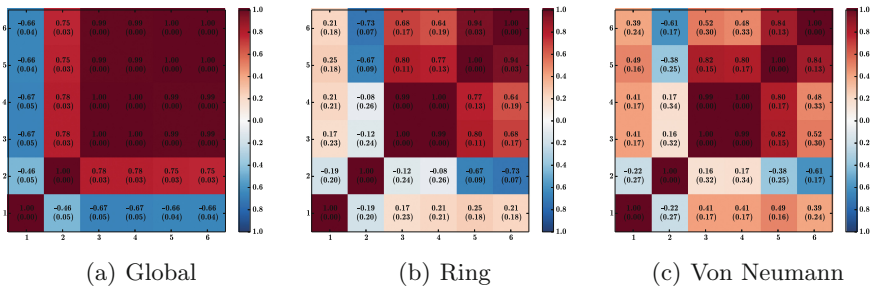

(a) Global   (b) Ring   (c) Von Neumann

Fig. 1. Correlations between Swarm Diversity Metrics using the Pearson Correlation Coefficient for Global (a), Ring (b) and (c) Von Neumann topologies. All simulations were repeated 30 times and the averages were considered.

The relationship between these swarm spatial diversity measures changes depending on the topology and some may not be appropriate for local topologies. The idea of a swarm center $\bar{x}$ might make sense for a global topology but not for a ring topology where more than one attractor might exist. Figure 1 depicts the correlations between these metrics regarding some main topologies and the only correlation which remains strong in all topologies are $\mathcal{D}_{\bar{x}}$ and $\mathcal{D}_{all}$.

This consistent disagreement between swarm diversity metrics based on particles' properties poses the question of what they are really measuring and if they can be applied to topologies other than the global topology. Indeed, although local topologies are less vulnerable to get trapped in local optima, the majority of research neglects them and focus their studies in the global topology only.

2.2 The Influence Graph

In the simple version of the PSO, at a given iteration t, each particle i solely gets information from its best neighbor $n_i(t)$. Since the swarm topology defines the neighborhood of each particle, this structure only bounds the range of such communication. On the other hand, the actual information exchange among the particles is described by the influence graph whose elements are defined as:

$$I_{t_{ij}} = \begin{cases} 1, & \text{if} \quad n_i(t) = j \quad \text{or} \quad n_j(t) = i, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

This graph represents the structure of the particles that shared information between them at a given iteration t and differs from the swarm topology (aka communication graph) which is just a static graph determining the neighborhood of each particle [10,11]. I_t is a snapshot of the swarm communication at iteration t, thus the interactions occurred in past iterations are not present in this graph. The information exchanges between the particles during the whole algorithm execution until iteration t can be conveniently defined as:

$$I_t^w = \sum_{i=1}^{t} I_i. \tag{8}$$

The result of this sum is a weighted graph in which the weights of the edges $I_{t_{ij}}^w$ are equal to the number of the times two particles i and j exchanged information during the algorithm execution [9]. The influence of each particle on each other during the whole history of the swarm until t is represented in I_t^w. In order to understand a more recent history of the swarm, an useful weighted influence graph at iteration t within a time window t_w is defined as follows:

$$I_t^{t_w} = \sum_{i=t-t_w+1}^{t} I_i, \quad \text{with} \quad t \geq t_w > 0. \tag{9}$$

In other words, $I_t^{t_w}$ is the communication structure of particles that communicated among themselves at most t_w iterations before the iteration t. The value
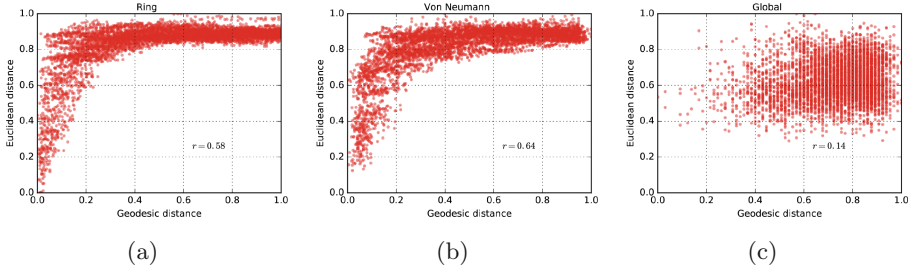
Fig. 2. The monitoring of the exchange of information in the swarm is enough to assess properties of the particles. The more distant two particles are in the influence graph, the farther they are from each other in the search space. The plot is the 1000th iteration of a run of the PSO algorithm optimizing the F_6 function [15].

of t_w changes the analysis of $I_t^{t_w}$ in such way that the lower t_w is, the shorter is the social memory being analyzed.

The influence graph provides insightful information about the intricate behavior of the swarm [9]. Even with the lack of details from the particles, the analysis of this graph allows one to examine the relationship between them. A simple supporting example of this is the distance between the particles in the search space. Figure 2 depicts the correlation between the geodesic distance among particles in the influence graph and the Euclidean distance between the particles in the search space. For a swarm with von Neumann or ring topologies, the more distant two particles are in the influence graph, the more certain we are that the two particles are far away of each other in the search space. This pattern is not as evident in the global topology due to the fact that the swarm, in this case, is condensed somewhere in the search space and the particles are nearer to each other than in the other topologies.

In fact, some analyses have been carried out with regards to the well-connected components in the influence graph [9]. Figure 3 exemplifies this kind of analysis, the edges are removed gradually according their weight in such way that components start to appear during the process. The speed that these components emerge as the destruction of the graph occurs is related to the search mode present in the swarm. For instance, an exploration mode is characterized by a slow increase in the number of components due to the different information flows present in the swarm [9]. On the other hand, this graph is rapidly destroyed in a swarm depending only on a small set of particles, a behavior related to an exploitation search mode.

3   Communication Diversity

The *structure* of the way information flows within the swarm can be assessed by the analysis of the well-connected components in the influence graph. The patterns found in this analysis are related to the diversity in the *communication* of the particles, *i.e.* the capacity of the swarm to have different information
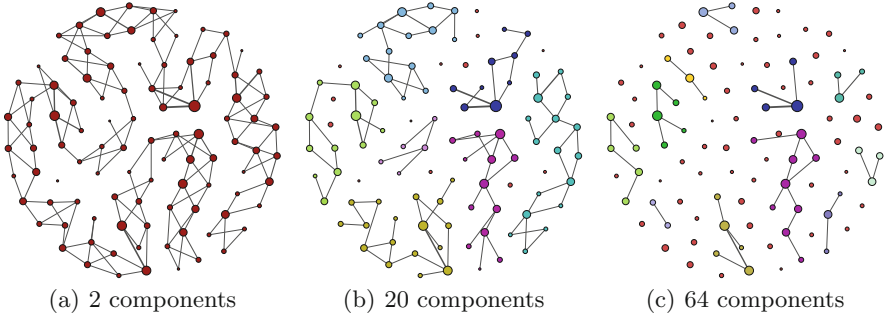
(a) 2 components   (b) 20 components   (c) 64 components

Fig. 3. The velocity components emerge while edges are removed from the influence graph is related to the search mode of the swarm. The edges are removed gradually depending on a threshold: below (*a*) 50 % of the highest edge weight, (*b*) 65 % and (*c*) 75 %. The different colors represent distinct components with more than one node. This is the 1000th iteration of a run of a PSO algorithm with the swarm using a von Neumann topology.

flows among its individuals. Figures 4(a)-(d) represent the communication diversity of the swarm. These plots depict the increasing number of components that emerge when edges are gradually removed from the influence graph with different time windows at the 1000th iteration of a run of a PSO algorithm on a ring/2-regular (Fig. 4(a)), von Neumann/4-regular (Fig. 4(b)), 30-regular (Fig. 4(c)) and global/100-regular (Fig. 4(d)) topologies. The gradual change in the number of components depends on the time window and represents the speed of the destruction occurring on the influence graph. A sharp behavior in the increase of the number of components, as seen in Figs. 4(c) and (d), is associated to lack of diversity in the communication, while a more gradual increase relates to more diversity, as depicted in Figs. 4(a) and (b).

The interactions between particles during the execution of a PSO algorithm might lead to variations in the communication diversity. For example, Fig. 4(e) shows how the number of components (*y*-axis) varies while applying different filters of weight removal on the influence graph with $t_w = 100$. For fixed topologies such as ring and von Neunmann, the number of components presents a consistently stationary behavior and reflects the static nature of these topologies. For a dynamic topology, the number of components could present an increasing or decreasing trend depending on how the communication structure changes towards the desired search mode. Besides, the variations in these curves along iterations (*x*-axis) suggest that an analysis of communication diversity needs to be performed also regarding time dimension. In order to carry out such analysis, we need to assess the *communication diversity*.

3.1   Assessing the Communication Diversity

Any column in the plots depicted in Figs. 4(a)-(d) represents the destruction of the influence graph with a certain t_w. A separate example of this curve is shown
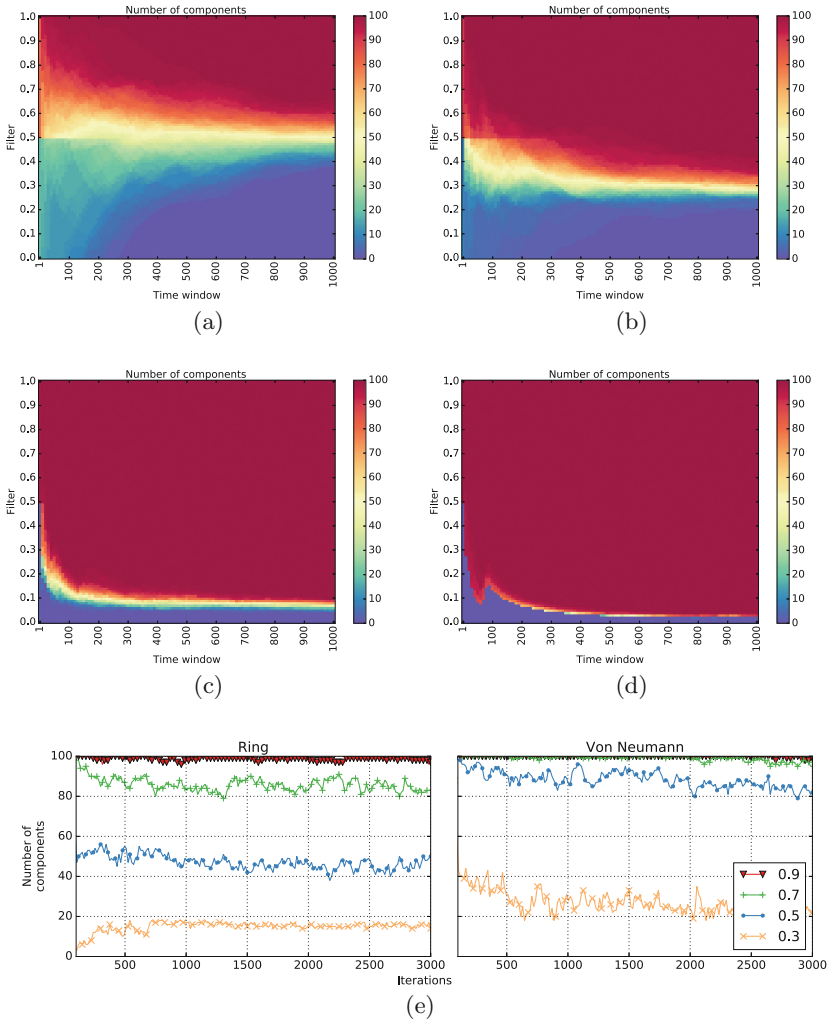
Fig. 4. The structure with which communication occurs within the swarm leads to distinct swarm behaviors. Each topology has its own communication diversity signature described by the number of components emerging (color intensity) as edges are removed (y-axis) of the influence graph with different time windows (x-axis). The edges are removed from the graph based on their weight and $2t_w$, *i.e.* the maximum weight possible in an influence graph with time window t_w. For a given filter, the number of components consistently varies within a well defined range through the iterations (e).

in Fig. 5 with $t_w = 100$ and $t_w = 1000$ for swarms with different topologies. Since such curves are always monotonically increasing, a simple way to have a characterization of them is to calculate the normalized area under the curve, namely $A_{t_w=t}$. Such procedure can be analogously performed with the snapshots
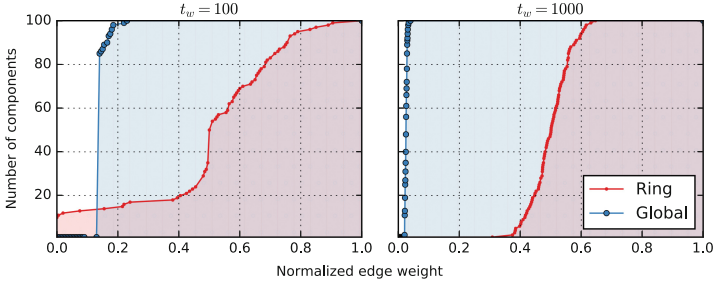
Fig. 5. The area under the curve of number of components that appear as the influence graph is destroyed can be used to compare swarm behaviors. The greater is the area, the less diverse is the communication of the swarm. For instance, a swarm with Global topology presents more exploitation than a swarm using the Ring topology.

of the communication diversity in Figs. 4(a)-(d) by taking into account all time windows. However, this approach can be computationally expensive given the great number of possible time windows. An approach to circumvent the computational cost is to use only a small set of time windows T in a such way that communication diversity (cd) can be assessed by:

$$cd = 1 - \frac{1}{|T||S|} \sum_{t \in T} A_{t_w = t}. \tag{10}$$

3.2 Communication Diversity and Stagnation

An implementation of the PSO algorithm with 100 particles was used to optimize the Ackley's F_6 *function* which is a shifted, single-group m-rotated, m-nonseparable and has many local minima. In all experiments, the number of dimensions was set to 1000 and m to 50, $c_1 = 2.05$, $c_2 = 2.05$, guaranteeing the algorithm to converge [3].

The experiments were performed with the swarm using connected k-regular graphs (*i.e.* graphs in which each node has k neighbors) as communication topology with $k = 2, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100$ (special cases are: $k = 2$, ring topology; $k = 4$, von Neumann topology; and $k = 100$, global topology). The PSO was executed 30 times for each topology and the influence graph was retrieved from the information exchange between the particles in each iteration of each execution. In order to analyze stagnation, we defined that the swarm stagnated at iteration t if the global best fitness did not improve by at least ρ between consecutive iterations until $t + \delta$ with $\rho = 1.02$ and $\delta = 500$.

Our results demonstrated the capability of *communication diversity* to explain the stagnation phenomenon for different communication topologies of Particle Swarm Optimizers using an unified approach based on graph component analysis. For instance, *communication diversity* (Fig. 6(b)) appear to firmly constrain the iteration PSO stagnates (Fig. 6(a)). In fact, by closely analyzing
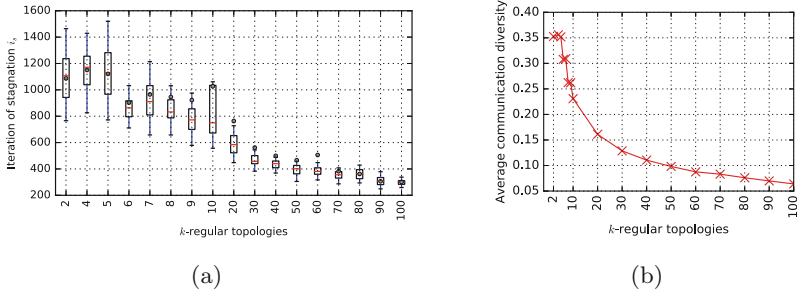
(a)   (b)

Fig. 6. The stagnation occurs earlier in swarms with k-regular topology as k increases (a), a tendency associated to the decrease of communication diversity (b). The boxplot (a) contains the value of the iteration the swarm stagnated in each of the 30 runs of the PSO algorithm, and the curve (b) depicts the average communication diversity until the iteration of stagnation.
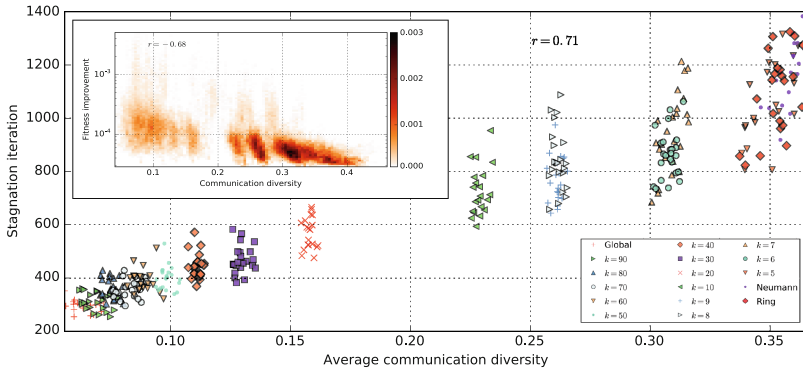

Fig. 7. The communication diversity of the swarm is associated to the iteration that stagnation occurs in the swarm. Each topology presents its range of communication diversity with some topologies sharing such range as well as iteration of stagnation. Each point is an execution of the 30 for each topology considered. Inset: Communication diversity is associated to the speed the global fitness f of the swarm changes between iterations t and $t+1$. The plot contains all the iterations until the swarm gets stagnated from 30 executions of PSO for each topology considered. The fitness improvement here is defined as $\frac{f_t - f_{t+1}}{f_{t+1}}$.

each execution (Fig. 7), different topologies can be clustered based on their *communication diversity*. These clusters captures the signatures associated with the communication diversity of each topology which plausibly explains, for instance, why the ring topology tends to consistently stagnate later when compared with the global topology. Surprisingly, *communication diversity* also relates to fitness improvement with a higher correlation (Fig. 7, inset) when compared with the other swarm spatial diversity measures albeit it only takes into account the

component analysis of the *influence graph* and disregards properties related to individual particles such as fitness, position or velocity.

The results suggest that swarm stagnation may be more related to how information flows through the swarm social structure than to the way specific particle properties are spread across the search space. For instance, simply restarting the swarm in order to highly disperse the particles might make sense considering the viewpoint of swarm spatial diversity but it does not make sense under the perspective of *communication diversity* for which this restarted swarm would be stagnated after the restart and still fated to consistently stagnate. Nevertheless, this also need to be confirmed in other problems and PSO variants.

4   Conclusions and Future Works

Early stagnation poses a barrier for Particle Swarm Optimizers and has only been tackled based on the assessment of the swarm spatial diversity. We advocate that our *communication diversity* based on component graph analysis of the *influence graph* is a good step towards an unified approach that can potentially better characterize early stagnation for different topologies.

Therefore, the design of stagnation prevention techniques should start concerning about *communication diversity* since it accounts for the social interactions among the individuals and not just properties associated with individual particles as the swarm spatial diversity. Approaches to control the communication diversity may be a fruitful way to control the execution of swarms in PSO and maybe other optimization approaches where an influence graph can be defined.

For future works, we will assess our *communication diversity* using other PSO variants in a panel of different problems, design a PSO variant with a stagnation prevention mechanism based on our *communication diversity* and extend the influence graph to other algorithms such as Ant Colony Optimization (ACO).

References

1. Bratton, D., Kennedy, J.: Defining a standard for particle swarm optimization. In: 2007 Swarm Intelligence Symposium, SIS 2007, pp. 120–127. IEEE, April 2007
2. Cheng, S., Shi, Y.: Diversity control in particle swarm optimization. In: 2011 IEEE Symposium on Swarm Intelligence (SIS), pp. 1–9, April 2011
3. Clerc, M., Kennedy, J.: The particle swarm - explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. **6**(1), 58–73 (2002)
4. Kennedy, J., Eberhart, R.: Particle swarm optimization, vol. 4, pp. 1942–1948 (1995)

5. Kennedy, J., Mendes, R.: Population structure and particle swarm performance. In: Proceedings of the 2002 Congress on Evolutionary Computation, CEC 2002, vol. 2, pp. 1671–1676 (2002)
6. Kennedy, J., Eberhart, R.C.: Swarm Intelligence. Morgan Kaufmann Publishers Inc., San Francisco (2001)
7. Krink, T., Vesterstrom, J., Riget, J.: Particle swarm optimisation with spatial particle extension. In: Proceedings of the World on Congress on Computational Intelligence, vol. 2, pp. 1474–1479 (2002)
8. Mendes, R., Kennedy, J., Neves, J.: The fully informed particle swarm: simpler, maybe better. IEEE Trans. Evol. Comput. **8**(3), 204–210 (2004)
9. Oliveira, M., Bastos-Filho, C.J.A., Menezes, R.: Towards a network-based approach to analyze particle swarm optimizers. In: 2014 IEEE Symposium on Swarm Intelligence (SIS), pp. 1–8, December 2014
10. Oliveira, M., Bastos-Filho, C.J.A., Menezes, R.: Using network science to assess particle swarm optimizers. Soc. Netw. Anal. Min. **5**(1), 1–13 (2015)
11. Oliveira-Júnior, M.A.C., Bastos-Filho, C.J.A., Menezes, R.: Assessing particle swarm optimizers using network science metrics. In: Ghoshal, G., Poncela-Casasnovas, J., Tolksdorf, R. (eds.) Complex Networks IV. SCI, vol. 476, pp. 173–184. Springer, Heidelberg (2013)
12. Olorunda, O., Engelbrecht, A.P.: Measuring exploration/exploitation in particle swarms using swarm diversity. In: 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence), pp. 1128–1134. IEEE, June 2008
13. Pontes, M.R., Neto, F.B.L., Bastos-Filho, C.J.: Adaptive clan particle swarm optimization. In: 2011 IEEE Symposium on Swarm Intelligence (SIS), pp. 1–6. IEEE (2011)
14. Shi, Y., Eberhart, R.: Monitoring of particle swarm optimization. Front. Comput. Sci. China **3**(1), 31–37 (2009)
15. Tang, K., Li, X., Suganthan, P.N., Yang, Z., Weise, T.: Benchmark functions for the CEC 2010 special session and competition on large-scale global optimization. Technical report, University of Science and Technology of China (USTC), School of Computer Science and Technology, Nature Inspired Computation and Applications Laboratory (NICAL), China
16. Zhan, Z.H., Zhang, J., Li, Y., Chung, H.H.: Adaptive particle swarm optimization. IEEE Trans. Syst. Man Cybern. Part B Cybern. **39**(6), 1362–1381 (2009)
17. Zhang, Q.L., Li, X., Tran, Q.A.: A modified particle swarm optimization algorithm. In: Proceedings of 2005 International Conference on Machine Learning and Cybernetics, vol. 5, pp. 2993–2995, August 2005