Communication, Leadership, Publicity and Group Formation in Particle Swarms*

Riccardo Poli[1], William B. Langdon[1], Paul Marrow[2], Jim Kennedy[3], Maurice Clerc[4], Dan Bratton[5], and Nick Holden[6]

[1] Department of Computer Science, University of Essex, UK
[2] BT Pervasive ICT Research Centre, Adastral Park, Ipswich, UK
[3] US Bureau of Labor Statistics, Washington DC, USA
[4] Independent Consultant, Groisy, France
[5] Department of Computing, Goldsmiths College, University of London, UK
[6] Computing Laboratory, University of Kent, UK

Abstract. We look at how the structure of social networks and the nature of social interactions affect the behaviour of Particle Swarms Optimisers. To this end, we propose a general model of communication and consensus which focuses on the effects of social interactions putting the details of the dynamics and the optimum seeking behaviour of PSOs into the background.

1 Introduction

In the standard Particle Swarms Optimiser, there are three features that bias the particle to look in a better place: 1) the particle remembers its best position, 2) it identifies its best neighbour, and 3) it knows that neighbour's best position so far. However, not all three forms are needed for good performance. For example, in Mendes' Fully Informed Particle Swarm (FIPS) model [1–3], the first two are absent. However, the topology of the social network of a PSO is considered a critical factor in determining performance.

Researchers have investigated how different topologies for the social network affect performance [1–8]. [4]. For example, it has been reported that with unimodal problems a *gbest* topology provides better performance, while the *lbest* PSO topology performs well on multimodal functions. Also, we know that with an appropriate topology, FIPS performs significantly better than the standard best-neighbour PSO on an array of test functions.

Albeit several lessons have been learnt in previous research on PSO social networks, the focus has mostly been on PSO performance rather than behaviour. However, it is clear behaviour is what matters, since performance on any particular fitness function is the result of coupling the features of that function with the natural behaviour of that particular PSO. If the mode of exploration fits the features of a problem, we expect good performance and vice versa.

* Work supported by EPSRC XPS grant GR/T11234/01.

Although PSOs are inspired by natural swarms, shoals, flocks, etc., the social network in a PSO has some important differences from its natural counterparts. In particular, a particle is fully aware of what happens in its neighbourhood. E.g., if at one iteration a particle in a neighbourhood achieves an improvement in its personal best, by next iteration all the other particles in that neighbourhood will be influenced by this change of state. It is as if the new local or global leader broadcasted its new state to the whole neighbourhood in one go. In nature this is not possible. Very often one can acquire information only by one-to-one interactions with another individual in the social network. So, the propagation of information is a stochastic diffusive process, rather than a deterministic broadcast.

Another feature of a PSO, which is not very common in nature, is that the best individual in the population, as long as it remains the swarm best, is unaffected by the others. In a shoal of fish, an individual may know about the presence of food or of a predator and act accordingly independently of other individuals. This individual will tend to act as a leader and will be followed by other fish. However, if, for whatever reasons, the rest of the shoal does not follow, it stops acting as a leader and rejoins the shoal.

Most researchers consider the social network as the communication medium through which information is exchanged. An important question is what exactly we mean by that. Perhaps a way to understand this is to investigate the properties of social networks as systems for reaching consensus. Ultimately the social network in a PSO is expected to get all of the particles to “agree” on where the swarm will search. However, the process and stages through which a global consensus is reached may be very different in PSOs using different communication topologies. For example, initially, until some form of consensus emerges, a population will act as a set of independent individuals. In systems where only localised interactions are possible (e.g., in a *lbest*-type of PSO) consensus must emerge firstly on a local basis and then progressively at larger and larger scales. However, when groups of individuals reach some local consensus, the group may start showing emergent properties. To understand the behaviour of the system as a whole it then becomes important to understand in which ways local groups (as opposed to single individuals) interact. This is particularly important in systems and landscapes where different parts of the population can reach different consensual decisions. If two such domains of agreement can come into contact, can they both persist? If not, will the consensus reached in one domain invade the other? Or, will the interaction produce a domain of a third type?

Another unnatural feature of PSOs, is that, once provided with a fitness function, they are closed systems. I.e. no external signals affect a PSO. In natural populations and many other systems there are always external influences. For example, in a social network of human customers or voters, individuals will influence each other and over time local or global consensus may emerge. However, the process of consensus formation and its final outcomes may be influenced by external factors, such as advertisements or broadcast party messages. These factors have been considered in agent-based models of customer relationship management where interactions between the social network of agents (customers) influence the response to marketing and other business activities (e.g., see [5]).

Naturally, it would be possible to modify PSOs to incorporate these three features: non-deterministic communication, democratic (as opposed to dictatorial) leaders and external influences. The question would then be, in what way would the search *behaviour* of a PSO be affected?

Rather than implementing PSO variants and testing them as optimisers for some set of functions (which we may do in future studies), we prefer to abstract away from the particular details of the dynamics of a PSO and to model social networks as consensus machines. In doing so we introduce approximations. So, in a sense, our model will not represent exactly any particular PSO. However, in return for these approximations we will gain a much deeper understanding of the social dynamics in particle swarms for different choices of interaction mechanisms and external influences.

The paper is organised as follows. In Section 2 we describe how we can abstract PSOs with very simple models which put many detailed aspects of the particle dynamics into the background. In Section 3 we consider the elitist communication strategies used in current PSOs and model them within a more general framework, which allows us to evaluate other, less-elitist strategies. These models are executable and so we were able to test them. Section 4 describes the parameters and fitness functions used in our experiments, while Section 5 reports on some key results. These are discussed in Section 6.

2 Abstracting PSO States and State Transitions

Let us consider a standard PSO. For as long as a particle's best and a neighbourhood best do not change, the particle will swarm around the point where the forces are zero. That is, around

$$x_i^* = \frac{x_{s_i} R_{\max_1} + x_{p_i} R_{\max_2}}{R_{\max_1} + R_{\max_2}}$$

where x_{s_i} is the i^{th} component of the best point visited by the neighbours of the current particle, x_{p_i} is the i^{th} component of its personal best, and $R_{\max_1}$ and $R_{\max_2}$ are the upper bounds of the intervals from which random numbers are drawn. This means that the points sampled by the PSO will be distributed about x_i^* until a new personal best or a new neighbourhood best are found. Because of this, on average a new personal best will tend to be found somewhere in between x_s and the old personal best. If the personal best gets closer to x_s, the new sampling distribution will progress even more towards x_s.

Naturally, when the distance between x^* and x_s reduces, the variance of the sampling distribution also reduces. This is effectively a step-size adjustment, which may be very important for good optimisation performance. The effect of this adjustment is that the probability of sampling the region "between" x_s and x_p remains high despite this region shrinking. If we took the centre of the sampling distribution x^* (rather than x_p) to represent the state of a particle, and approximate the whole swarm as a set of stochastic oscillators, each centred at its own x^*. We could then approximately model the dynamics of x^* by imagining that, on average, when a particle interact socially with another particle in its neighbourhood a sort of stochastic interpolation process between the particles' states takes place.[7]

A natural modelling choice for PSOs is to consider states as continuous variables which change with an interpolation rule such as

$$x^*_{new} = \alpha x^*_s + (1 - \alpha) x^*_{old} \tag{1}$$

where x^*_{old} is the current state of a particle, x^*_s is the state of the particle with which interaction takes place (e.g., the neighbourhood best), and α is a random variable uniformly distributed in $[0, 1]$. These choices seem particularly suitable for smooth monotonic landscapes and for PSOs where velocities are limited (e.g., when ψ_1 and ψ_2 are small or friction is high). In highly multimodal landscapes each particle's best and, consequently x^*, may change less continuously.

To explore what happens at this other extreme of the behaviour, we idealise the multi-modal case and imagine that only a finite number, n, of different states x^* are possible. State transitions can then be represented using a Markov chain, where the entries of state transition matrix represent the probability of x^* moving from one state to another. The transition probabilities would have to be such as to effectively implement a stochastic interpolation process between x^* and x_s like the one in Equation 1.

If n is a power of two, a simple realisation of this is to represent states in binary form. So, a generic state x could be represented as $x_1 x_2 \cdots x_\ell$ (with $\ell = \log_2 n$). The interpolation process between states could then be implemented by the replacement of one or more random bits in x^* with corresponding bits from x_s. That is, for discrete states we could use the update rule: $x^*_{new_i} = x^*_{old_i}$ with probability β and $x^*_{new_i} = x^*_{s_i}$ with probability $1-\beta$, where β is a constant.

Irrespective of the model used, we can see that state update rules have the form $x^*_{new} = \chi(x^*_s, x^*_{old})$ where χ is a stochastic function.

In a standard PSO, and also in many natural and artificial systems, an individual's choice as to whom to interact with depends on quality, desirability, or, more generally, on whatever is an appropriate measure of success for an individual. We will just call this *fitness*. Without any information about the fitness function it would be impossible to say much about the behaviour of a system. So, even in our effort to hide the finer details of the dynamics of PSOs, we will need to keep a notion of fitness. In particular, we will imagine that each state, x, has an associated fitness value f_x. Note that this may be distinct from the value taken by the fitness function, f, in x (seen as a point in the search space), $f(x)$. For example, for standard PSOs, we could take this to be the value of fitness associated with the current particle best, i.e. $f_x = f(x_p)$, or the expected fitness value for a sampling distribution centred at x^*. In a maximisation

[7] This is an approximation which is expected to be reasonably accurate if we look at states at a suitably coarse time scale, and, of course, as long as x_s remains constant.

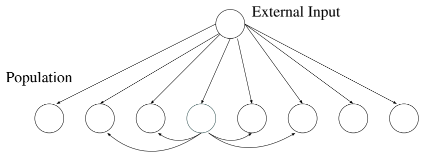

Fig. 1. Extended *lbest* social network topology.

problem, with any reasonable definition of f_x, in a PSO we would expect to see $f_{x_s} \geq f_{x^*_{new}} \geq f_{x^*_{old}}$.

If we use integer states then we can represent the state-fitness function f_x as a table. For example, if the search space was one-dimensional then the table would look like the following

x	0	1	$\cdots$	$n-1$
f_x	f_0	f_1	$\cdots$	f_{n-1}

In order to keep things simple, for our real-valued state representation we will define f_x explicitly only over a discrete lattice of points in the search space and then use linear interpolation to construct fitness values elsewhere. This allows us to still represent f_x as a table, albeit at the cost of a reduced generality.[8]

3 Particle communication as it is and as it could be

We are interested in looking at current forms of inter-particle communication in PSOs as part of a bigger family of social interactions. To allow the study of the effects of external sources of information, noise or publicity we extended the classical *lbest* ring topology as shown in Figure 1. We imagine that the external input is simply another individual which is in the neighbourhood of every particle and which has a pre-defined and constant state x_{ext}. Naturally, we don't necessarily want the external input to act at all times and for all individuals at the same time. So we model the effect of the external input using a probabilistic update rule: $x^*_{new} = \chi(x_s, x^*_{old})$ with probability γ, and $x^*_{new} = \chi(x_{ext}, x^*_{old})$ with probability $1 - \gamma$, where γ is a constant controlling the intensity of the action of the external input. When $\gamma = 0$ the system is again a closed system.

As we mentioned in Section 1, we want to study the effects of reducing the ability of individuals to perceive the state of their neighbourhood. This implies that it is possible for them to decide to have social interactions with individuals other than the neighbourhood best. We do this by defining a stochastic social-selection function σ which, given the states x_1, x_2, etc. of the individuals in

[8] PSOs and other rank based optimisers behave identically on any pair of isomorphic landscapes f and g such that for every pair of points x and y, $f(x) > f(y)$ if and only if $g(x) > g(y)$. So, fine details on the shapes of landscapes are often not very important.

the neighbourhood returns one of such states. For a standard PSO we have $\sigma(x_1, x_2, \cdots) = \text{argmax}_{x\in\{x_1,x_2,\cdots\}} f(x)$. Naturally, this form of selection function guarantees that, if an individual is the neighbourhood best, the state of that individual will always be returned by σ. Since the state update functions χ defined above are such that $\chi(x, x) = x$, in these circumstances

$$x^*_{new} = \chi(\sigma(\cdots, x^*_{old}, \cdots), x^*_{old}) = \chi(x^*_{old}, x^*_{old}) = x^*_{old}$$

which guarantees that neighbourhood-best individuals are not affected by state updates.

To explore what happens at the other extreme of the spectrum, while still favouring social interactions with above average fitness individuals, we introduce another version of σ where the value returned is the result of a binary tournament. That is, we draw (with repetition) two random states from the set of neighbouring states $\{x_1, x_2, \cdots\}$. Out of these two, we then return the state with higher fitness. That is

$$\sigma(x_1, x_2, \cdots) = \begin{cases} x' & \text{if } f(x') > f(x'') \\ x'' & \text{otherwise} \end{cases} \tag{2}$$

where x' and x'' are two randomly (and uniformly) chosen states from the set $\{x_1, x_2, \cdots\}$. With this social-selection rule the neighbourhood best individual can still be required to interact with some individuals other than itself. This means that its state may change and it may no longer remain a leader forever. Unlike the standard PSO, this leadership change can happen even if an individual with a better state has not been found in the neighbourhood.

4 Experimental Setup

The models described in the previous section are executable models. Once a fitness table f_x, the size of the population and the structure of the social network are defined, it is possible to iterate the state update and inter-particle interaction rules for as many iterations (generations) as desired. The objective of running these models, however, is not to see how well the population can locate global optima. Rather, we want to see the different qualitative behaviours the different types of social interactions and states can provide.

For this study we decided to keep the search space one-dimensional. Both for the integer state representation and the real-valued representation we used numbers in the range [0,7]. For the integer representation $n = 8$, so we used $\ell = 3$ bits. We used populations of $P = 80$ individuals (we chose a relatively large population because this slows down all transients, making it possible to better see the phenomena we are interested in). Runs lasted for 200 generations. In the initial generation, binary individuals were given random states. (Since there are only 8 possible states, with a population of 80, in virtually all runs, all states had non-zero frequency). In the case of the real-valued representation,to ensure a wide spread, we randomly initialised individuals by drawing states *without* *replacement* from the set $\left\{\frac{7\times i}{79}\right\}_{i=0}^{79} = \{0, 7/79, 14/79, \ldots, 7\}$. With our choice of fitness functions (see below), these initialisation strategies virtually guarantee that the global optimum is always represented in the initial population. This puts the search for optima in the background, allowing us to focus on the dynamics of information flows and on consensus/group formation and stability.

We considered *lbest*-type (ring) topologies for the social network. However, we tested values for the neighbourhood radius r in the set $\{1, 2, 3, 4, 5, 10, 20, 40\}$, thereby going from the most extreme form of local interaction to a *gbest*-type of topology (with $P = 80$, an *lbest* topology with $r = 40$ is a fully connected network).

We tested applying the external input at probabilities of $\gamma \in \{0, 0.01, 0.03, 0.05, 0.1, 0.2, 0.5, 1\}$ (per individual). When $\gamma > 0$, in both binary and real-valued state representations, we tested two different values (2 and 7) for x_{ext}. We tested both the deterministic social-selection strategy (typical of PSOs) and the probabilistic one based on tournaments. In the binary representation, social interaction occur by randomly exchanging bits with a probability of $\beta = 0.5$ (per bit).

For all settings we tested the six fitness functions shown in Table 1.

Combined together the experimental settings described above produce over 3,000 conditions. In each condition, we gathered statistics over 200 independent runs of the model, for a total of over 600,000 runs. Due to space limitations, we are able to report only a tiny subset of our experimental results here. However, a fuller set of results is provided online in [6].

5 Results

For each of the settings described in the previous section, we did both single runs and multiple runs to assemble statistically reliable state histograms. The single runs are represented by 3–D plots of the state of each individual in the population in each generation. (For space limitations here we do not show any such runs, but in [6] we plots 144 of them.) The state histograms represent the average behaviour of the system over 200 independent runs. These are also represented by 3–D plots, but this time they represent how the proportion of individuals in each particular state changed over time (so the topological organisation of the individuals is not represented). In these plots the x axis represent the generation number, the y axis the state and the z axis the proportion of individuals in that state at that generation. For binary states only 8 different states are possible and so we collected statistics for all of them. For maximum comparability, for the continuous representation, we divided the state space into 8 bins, centred at 0, 1, 2, 3, 4, 5, 6 and 7. States were associated to the nearest bin. (As a consequence, in the initial generation bins 0 and 7 have only half the number of samples of other bins.)

As shown in Figure 2, with a real-valued representation, deterministic communication and in the absence of exogenous inputs, the model behaves (as expected) like a PSO. The only difference between the small neighbourhood ($r = 1$)

Table 1. Fitness functions used in this study.

	Fitness values								
Function	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	Description
flat	1	1	1	1	1	1	1	1	a base line case that reveals the natural biases associated with the social interactions
linear	1	2	3	4	5	6	7	8	a function where different states are associated with different fitness, and where individuals that are neighbours *in state space* have similar fitness
twopeaks	8	4	4	4	4	4	4	8	a function with two symmetric peaks at opposite extremes of the state space embedded in a flat landscape that is useful to study the stability of consensus domains
nondeceptive	8	2	2	2	2	2	2	7	a function almost identical to *twopeak*, but where one peak is actually a local optimum, which provides a comparative basis for other functions
deceptive (binary)	8	2	2	4	2	4	4	7	a function with the same optima as *nondeceptive*, but where the neighbours of the optima in state space have different fitness, which allows us to show why the consensus reached by the population may be counterintuitive
deceptive (float)	8	2	2	2	4	4	4	7	see binary deceptive fitness function
trap	8	1	2	3	4	5	6	7	a variant of *linear* where the global optimum is in the area of the search space with the lowest fitness and all gradient information points towards the local optimum, thereby making the problem very deceptive

case, shown in Figure 2, and other cases is the speed at which a steady state is reached. Note: there are dynamics in the system even on a flat landscape. In particular, the interactions produce an implicit bias towards the average state value (3.5) (Figure 2 top left). The model produces a distribution which is very similar to the bell-shaped sampling distribution obtained in real PSOs when x_s and x_p are kept constant [7], which corroborates the approach. Note that this bias can be masked by the presence of fitness gradients, but cannot be removed.

In addition, we can see that local optima have an effect, with initially a part of the population (except for a fully connected swarm, $r = 40$) forming a consensus towards exploring them (Figure 2 middle right, bottom left and right). This slows down convergence to the global optimum. This effect, however, does not simply depend on local optima, but is also markedly influenced by the fitness of neighbouring individuals. This is a form of *deception*. The deception is only partial since, with real-valued states and deterministic communication, eventually all individuals settle for the highest fitness state. If we change the social-selection strategy to tournaments (cf. Equation 2), however, we see that

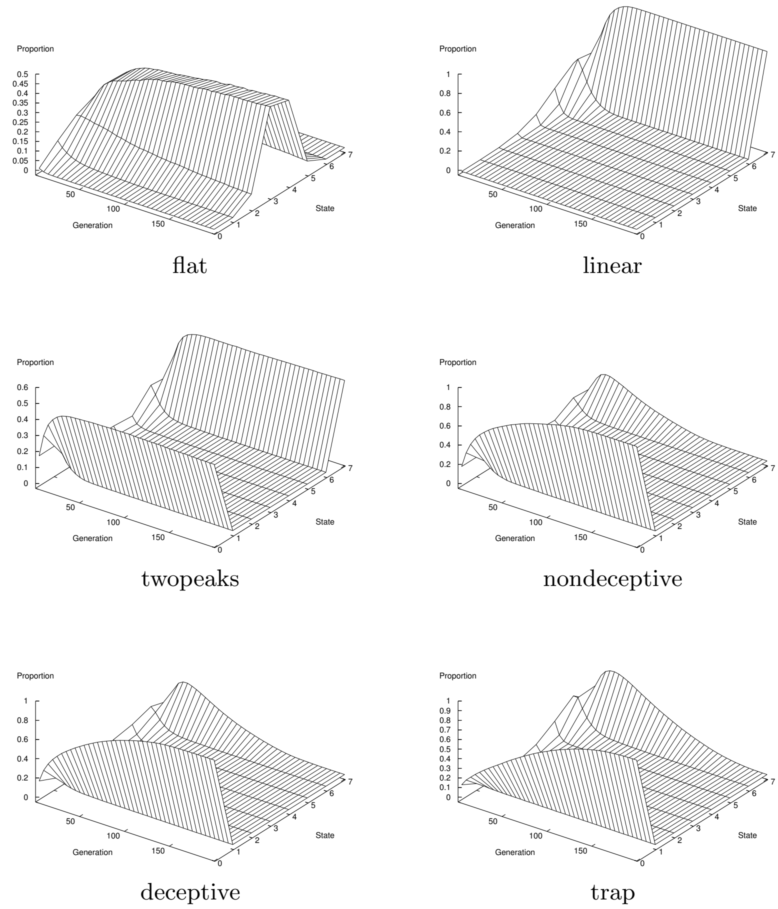

Fig. 2. State evolution histograms for the real-valued model with deterministic communication, $r = 1$ and no external input, for the 6 fitness functions used in our study.

the deceptive pressure can become predominant and force the population towards a suboptimal state, as shown in Figure 3. Even with bigger neighbourhoods the swarm will be deceived in these conditions. The only difference is that with more global communication the algorithm settles for some neighbour of the local optimum, due to the implicit bias of the interaction mechanism discussed above (see Figure 4). Interestingly, a form of deception has been recently reported in real PSOs [8, 9], which again corroborates our model.

The injection of an external input $x_{ext} = 2$ with a very low probability ($\gamma = 0.03$) has a clear impact on behaviour. Except for *flat*, for the functions we considered $x = 2$ is not a desirable state. So, this input can be considered as publicity for a suboptimal product (state). We can see the effects of exogenous inputs by considering the flat fitness case. What happens in this case, is that the

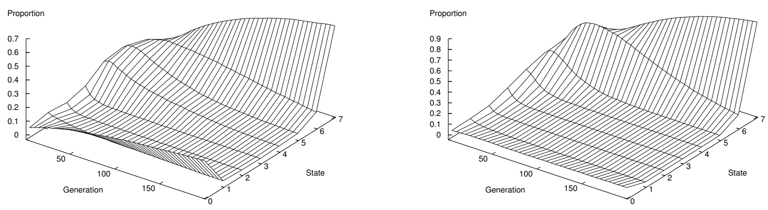

Fig. 3. State evolution histograms for the real-valued model with *probabilistic* communication, $r = 1$ and no external input, for the *deceptive* (left) and *trap* (right) functions.

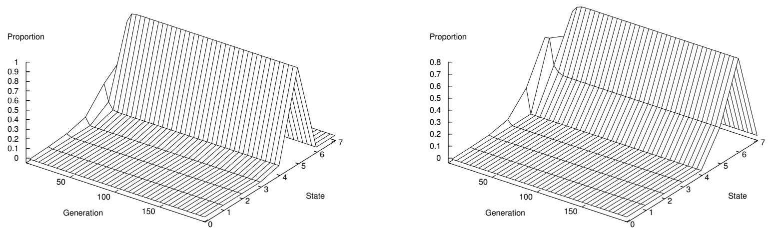

Fig. 4. As Figure 3 but for $r = 40$.

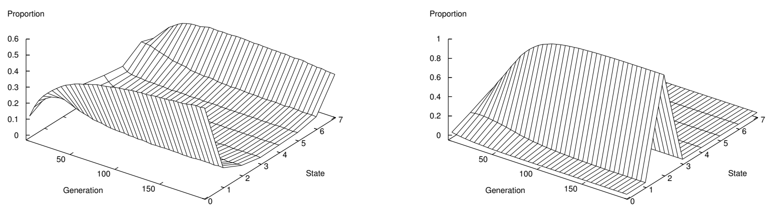

Fig. 5. State evolution histograms for the real-valued model with deterministic (left) and probabilistic (right) communication ($r = 1$, $x_{ext} = 2$, $\gamma = 0.1$, and $f_x =$*twopeaks*).

population initially moves towards the natural fixed-point state (3.5), but later on, with the reiterated introduction of spurious states, eventually converges to state 2. This bias cannot be completely eliminated in non-flat landscapes, and it can cause, for example, the breaking of symmetries, as shown in Figure 5 left, or even the convergence to a low-fitness state, as shown in Figure 5 right.

6 Discussion

We have proposed a simplified but general model of communication and consensus dynamics in PSOs. The model was specifically designed to look at how the structure of the social network and the nature of the social interactions affect the behaviour of these systems. So, we made an effort to conceal as much as possible of the implementation details, of the dynamics and of the optimum seeking behaviour of PSOs.

Models are useful tools to understand systems, but, except for very simple systems, no model can tell everything there is to know. That is, every model will make it easy to answer certain questions, and hard or impossible to answer different questions. That is why it is important to build models starting from different points of view. For PSOs, nothing can replace dynamical system models of PSOs [10]. However, these models become immensely complex and are difficult to study unless one makes simplifications. It is not easy to imagine how one could, for example, incorporate different topologies, different fitness functions, etc. in such models and still be able to get qualitative answers, without approximations.

The objective of our model is not to replace the dynamical systems approach. It is to complement it. The model focuses on social interactions, not fine details of the dynamics. So, there will be surely many things that this model cannot capture. In return, however, the model makes it easier to ask and get answers to other questions.

With all models, they are only as good as the predictions and understanding they can produce. These need to be checked against empirical data for further corroboration. The model has already provided new insights into particle swarms. For example, it has highlighted how the swarm consensus can be deceived away from the optimal state and how exogenous sources of influence can break symmetries, modify natural search biases and even lead the PSO to completely ignore fitness. Some of its predictions match results obtained by other means. Other predictions will need to be checked by implementing and testing new PSOs, e.g., PSOs with non-deterministic communication and PSOs with external inputs.

The model includes the forms of communication currently implemented in PSOs, but it is significantly more general. As a result, we believe that this model and our results may be applicable to natural and artificial systems other than just particle swarms (e.g., social networks of customers). We will explore this in future research.

References

1. Mendes, R., Neves, J.: What makes a successful society? experiments with population topologies in particle swarms. In Bazzan, A.L., Labidi, S., eds.: Advances in Artificial Intelligence. XVII Brazilian Symposium on Artificial Intelligence - SBIA'04. Volume 3171 of Lecture Notes in Computer Science., So Lus, Maranho – Brazil, Springer (2004) 346–355
2. Mendes, R., Kennedy, J., Neves, J.: The fully informed particle swarm: Simpler, maybe better. IEEE Transactions of Evolutionary Computation **8**(3) (2004) 204–210
3. Mendes, R., Kennedy, J., Neves, J.: Avoiding the pitfalls of local optima: How topologies can save the day. In: Proceedings of the 12th Conference Intelligent Systems Application to Power Systems (ISAP2003), Lemnos, Greece, IEEE Computer Society (2003)
4. van den Bergh, F.: An Analysis of Particle Swarm Optimizers. PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa (2001)
5. Baxter, N., Collings, D., Adjali, I.: Agent-based modelling – intelligent customer relationship management. BT Technology Journal **21**(2) (2003) 126–132
6. Poli, R., Langdon, W.B., Marrow, P., Kennedy, J., Clerc, M., Bratton, D., Holden, N.: Communication, leadership, publicity and group formation in particle swarms. Technical Report CSM-453, Department of Computer Science, University of Essex (2006)
7. Kennedy, J.: Bare bones particle swarms. In: Proceedings of the IEEE Swarm Intelligence Symposium (SIS) 2003, Indianapolis, Indiana (2003) 80–87
8. Langdon, W.B., Poli, R., Holland, O., Krink, T.: Understanding particle swarm optimisation by evolving problem landscapes. In Gambardella, L.M., Arabshahi, P., Martinoli, A., eds.: Proceedings SIS 2005 IEEE Swarm Intelligence, Pasadena, California, USA, IEEE (2005) 30–37
9. Langdon, W.B., Poli, R.: Evolving problems to learn about particle swarm and other optimisers. In Corne, D., Michalewicz, Z., Dorigo, M., Eiben, G., Fogel, D., Fonseca, C., Greenwood, G., Chen, T.K., Raidl, G., Zalzala, A., Lucas, S., Paechter, B., Willies, J., Guervos, J.J.M., Eberbach, E., McKay, B., Channon, A., Tiwari, A., Volkert, L.G., Ashlock, D., Schoenauer, M., eds.: Proceedings of the 2005 IEEE Congress on Evolutionary Computation. Volume 1., Edinburgh, UK, IEEE Press (2005) 81–88
10. Clerc, M., Kennedy, J.: The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation **6**(1) (2002) 58–73