Effects of Scale-Free and Small-World Topologies on Binary Coded Self-Adaptive CEA

Mario Giacobini[1], Mike Preuss[2], and Marco Tomassini[1]

[1] Information Systems Department, University of Lausanne, Switzerland
mario.giacobini@unil.ch, marco.tomassini@unil.ch
[2] Systems Analysis Group, Computer Science Department, University of Dortmund, Germany
mike.preuss@uni-dortmund.de

Abstract. In this paper we investigate the properties of CEAs with populations structured as Watts–Strogatz small-world graphs and Albert–Barabási scale-free graphs as problem solvers, using several standard discrete optimization problems as a benchmark. The EA variants employed include self-adaptation of mutation rates. Results are compared with the corresponding classical panmictic EA showing that topology together with self-adaptation drastically influences the search.

1 Introduction

The standard population structure used in evolutionary algorithms (EAs) is the *panmictic* structure. In panmictic populations, also called mixing, any individual is equally likely to interact with any other individual. This setting is the most straightforward and many theoretical results have been obtained for it. However, since at least two decades, several researchers have suggested that EAs populations might have structures endowed with spatial features, like many natural populations (for recent reviews see [1, 2] and references therein). Empirical results suggest that using structured populations is often beneficial owing to better diversity maintenance, formation of niches, and lower selection pressures in the population favoring the slow spreading of solutions and relieving premature convergence and stagnation. The most popular models are the *island model* and the *cellular model*. In the island model the whole population is subdivided into several subpopulations each of which is panmictic. A standard EA runs in each subpopulation and, from time to time, a fraction of individuals migrate between islands. Although this model may offer some advantages over a single mixing population, it is still rather close to the latter.

Here we shall focus on cellular models instead, which are a more radical departure from the standard setting. What sets them apart is the fact that all the operators act locally, within a small pool of individuals. The customary cellular topology is the regular lattice. Cellular evolutionary algorithms (CEAs) on regular lattices, usually rings and two-dimensional grids, have been often used with good results and some of their theoretical properties are known (see [2]). However, there is no reason why cellular models should be limited to regular lattices. Other graph structures are possible, such as random graphs and *small-world* networks. These small-world networks are not regular nor completely random, and have recently attracted a lot of attention in many areas because

of their surprising topological properties [3, 4]. Random graphs and small-world networks have been recently studied from the point of view of the selection intensity in the population [5]. Random graphs are roughly equivalent to panmictic structures in behavior, at least for not too small probability of having an edge between two arbitrary vertices. The families of small-world graphs are potentially more interesting, as they can induce widely variable global selection pressures, depending on the value of some graph characteristic parameter [5]. A first investigation on the use of such structured populations for optimization problems has been proposed by Preuss and Lasarczyk [6].

In this paper we investigate the properties of CEAs with populations structured as Watts–Strogatz small-world graphs and Albert–Barabási scale-free graphs as problem solvers, using several standard discrete optimization problems as a benchmark. We should like to point out at the outset that it is not our intention to compete with the best heuristics for the problems. We do not use problem information, nor do we include any kind of local or enhanced search. Our goal is simply to compare these irregular population structures with regular lattices CEAs and the panmictic EA using the simplest settings and only few parameters. We are especially interested in answering the following questions:

- What is the influence of different node degree distributions on CEAs when the overall connectivity (number of connections) remains constant?
- Are scale-free topologies worthwhile alternatives to standard small-world ones? If so, for which problem types?
- When —if at all— does self-adaptation of mutation parameters provide an advantage over fixed mutation rates?

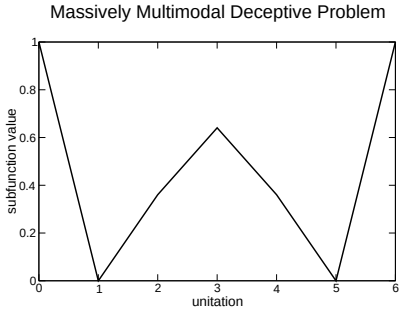
When dealing with evolutionary algorithms on binary represented problems, a sporadically suggested [7] and rarely used technique is the self-adaptation of mutation parameters. Although well established for continuous representations [8], its applicability is rather unclear for test problems typically approached with genetic algorithms. It is our hope that self-adaptation proves worthwhile for CEAs, especially in connection with small-world topologies.

2 Test Problems

In this section we present the set of problems chosen for this study. The benchmark is representative because it contains many different interesting features in optimization, such as epistasis, multimodality, deceptiveness, and problem generators. These are important ingredients in any work trying to evaluate algorithmic approaches with the objective of getting reliable results, as stated by Whitley et al. in [9].

We experiment with the massively multimodal deceptive problem (MMDP), a modified version of the multimodal problem generator P-PEAKS, error correcting code design (ECC), and the countsat problem (COUNTSAT). The choice of this set of problems is justified by both their difficulty and their application domains (combinatorial optimization, telecommunications, etc.). This gives us a fair level of confidence in the results, although no benchmark will ever be able to assert the superiority of a particular algorithm on all problems and problem instances [10]. The problems selected for this benchmark are briefly presented in the following paragraphs.

Massively Multimodal Deceptive Problem (MMDP). The MMDP is a problem that has been specifically designed to be difficult for an EA [11]. It is made up of k deceptive subproblems (s_i) of 6 bits each, whose value depends on the number of ones (*unitation*) a binary string has (see Figure 1). These subfunctions possess two global maxima and a deceptive attractor in the middle point.



unitation	subfunction value
0	1.000000
1	0.000000
2	0.360384
3	0.640576
4	0.360384
5	0.000000
6	1.000000

Fig. 1: *Basic deceptive bipolar function (s_i) for MMDP.*

In MMDP each subproblem s_i contributes to the fitness value according to its *unitation* (Figure 1). The global optimum has a value of k and it is attained when every subproblem is composed of zero or six ones. The number of local optima is quite large (22^k), while there are only 2^k global solutions. Therefore, the degree of multimodality is regulated by the k parameter. To avoid floor and ceiling effects (none or all EA are able to solve the problem) we use a moderately difficult instance with $k = 20$. Fitness is computed after after Eq. 1, utilizing subfunction s_i as depicted in Figure 1. Note that this problem is separable; its constituents could be optimized individually if its boundaries in the genome were known to the EA.

$$f_{MMDP}(\boldsymbol{s}) = \sum_{i=1}^{k} \mathit{fitness}(s_i) \tag{1}$$

Multimodal Problem Generator (wP-PEAKS). A problem generator is an easily parameterizable task which has a tunable degree of epistasis, thus permitting to derive instances with growing difficulty at will. With a problem generator we evaluate our algorithms on a high number of random problem instances. Since a different instance is solved each time the algorithm runs, the predictive power of the results for the problem class as a whole is increased.

The idea of P-PEAKS is to generate P random N-bit strings that represent the location of P peaks in search space. Using a small/large number of peaks results in weakly/strongly epistatic problems. In the original problem formulation [12], the fitness value of a string was the number of bits it had in common with the nearest peak in that space, divided by N. However, each peak represented a global optimum. We modified the problem by adding weights $w_i \in \mathbb{R}_+$ with only $w_1 = 1.0$ and $w_{[2...P]} < 1.0$,

thereby requiring the optimization algorithm to find the one peak bearing the global optimum instead of just any peak. It should be noted that doing so for one global and nine local peaks (as utilized in our experiments) —tested empirically— appears to be a lot harder than a standard P-PEAKS with $P = 100$.

$$f_{wP-PEAKS}(\boldsymbol{x}) = \frac{1}{N} \max_{1 \leq i \leq p} \{w_i \cdot N - HammingD(\boldsymbol{x}, Peak_i)\} \tag{2}$$

Error Correcting Code Design Problem (ECC). The ECC problem was presented in [13]. We will consider a three-tuple (n, M, d), where n is the length of each codeword (number of bits), M is the number of codewords, and d is the minimum Hamming distance between any pair of codewords. Our objective will be to find a code which has a value for d as large as possible (reflecting greater tolerance to noise and errors), given previously fixed values for n and M. The problem we have studied is a simplified version of that in [13]. In our case we search half of the codewords ($M/2$) that will compose the code, and the other half is made up by the complement of the codewords computed by the algorithm. The fitness function to be maximized is:

$$f_{ECC} = \frac{1}{\sum_{i=1}^{M} \sum_{j=1, i \neq j}^{M} d_{ij}^{-2}}, \tag{3}$$

where d_{ij} represents the Hamming distance between codewords i and j in the code C (made up of M codewords of length n). In the present paper, we consider an instance with $M = 24$ and $n = 12$, yielding optimum fitness of 0.0674 [14].

COUNTSAT Problem The COUNSAT problem has been proposed by Droste *et al.* [15] as an instance of the MAXSAT problem difficult to be solved by Evolutionary Algorithms. In COUNTSAT, the solution value is the number of clauses (among all the possible 3-variables Horn clauses) that are satisfied by an n-bit input string, where the binary value 0 and 1 are considered as a *false* and a *true* boolean value, respectively. It is easy to check that the optimum value is that of the solution with all the variables assigned to 1. Droste *et al.* have proved that the fitness of a tentative solution $\boldsymbol{x}$ can be easily computed using the following equation:

$$f_{\text{COUNTSAT}}(\boldsymbol{x}) = s + n(n-1)(n-2) - 2(n-2)\binom{s}{2} + 6\binom{s}{3}, \tag{4}$$

where s is the unitation of the solution $\boldsymbol{x}$ (i.e. the number of 1 entries in $\boldsymbol{x}$), and n is the length of $\boldsymbol{x}$. In this paper we will study an instance of $n = 20$ variables, with normalized optimum fitness of 1.0.

3 Small-World Graph Topologies

It has been shown in recent years that graphs occurring in many social, biological, and man-made systems are often neither completely regular, such as lattices, nor completely random [16]. They have instead what has been called a *small-world* topology, in which

nodes are highly clustered yet the path length between them is small. This behavior is due to the presence of *shortcuts* i.e., a few direct links between nodes that would otherwise be far removed. Following Watts' and Strogatz's discovery, Barabasi *et al.* [3] found that several important networks such as the World Wide Web, Internet, author citation networks, and metabolic networks among others, also have the small world property but their degree distribution function differs: they have more nodes of high degree that are likely in a random graph of the same size and edge density. These graphs have been called *scale-free* because the degree probability distribution function follows a power law. In the next sections we briefly describe how small-world and scale-free graphs can be constructed, more details can be found in [16, 3, 4].

3.1 The Watts–Strogatz Model

Although this model has been a real breakthrough in the technical sense when it appeared, today it is clear that it is not a good representation of real networks as it retains many features of the random graph model. In spite of this, the Watts–Strogatz model, because of its simplicity of construction and the richness of behavior, is still an interesting topology in artificial systems where there is no "natural" constraint on the type of connectivity.

According to Watts and Strogatz [16], a small-world graph can be constructed starting from a regular ring of nodes in which each node has k neighbors ($k \ll N$) by simply systematically going through successive nodes and "rewiring" a link with a certain probability β. When the edge is deleted, it is replaced with an edge to a randomly chosen node. If rewiring an edge would lead to a duplicate edge, it is left unchanged. This procedure will create a number of *shortcuts* that join distant parts of the lattice.

Shortcuts are the hallmark of small worlds. While the average path length[1] between nodes scales logarithmically in the number of nodes for a random graph, in Watts-Strogatz graphs it scales approximately linearly for low rewiring probability but goes down very quickly and tends to the random graph limit as β increases. This is due to the progressive appearance of shortcut edges between distant parts of the graph, which obviously contract the path lengths between many vertices. However, small worlds typically have a higher clustering coefficient[2] than random graphs. Small-world networks have a degree distribution $P(k)$ close to Poissonian.

3.2 The Barabási-Albert Model

Albert and Barabási were the first to realize that real networks grow incrementally and that their evolving topology is determined by the way in which new nodes are added to the network and proposed an extremely simple model based on these ideas [3]. At the beginning one starts with a small clique of m_0 nodes. At each successive time step a new node is added such that its $m \leq m_0$ edges link it to m nodes already in the

[1] The average path length L of a graph is the average value of all pair shortest paths.

[2] The clustering coefficient C of a node is a measure of the probability that two nodes that are its neighbors are also neighbors among themselves. The average $\langle C \rangle$ is the average of the Cs of all nodes in the graph.

graph. When choosing the nodes to which the new nodes connects, it is assumed that the probability π that a new node will be connected to node i depends on the degree k_i of i such that nodes that have already many links are more likely to be chosen over those that have few. This is called *preferential attachment* and is an effect that can be observed in several real networks. The probability π is given by:

$$\pi(k_i) = \frac{k_i}{\sum_j k_j},$$

where the sum is over all nodes already in the graph. The model evolves into a stationary scale-free network with power-law probability distribution for the vertex degree $P(k) \sim k^{-\gamma}$, with $\gamma \sim 3$.

4 Experiment

Focus. Investigate the effects of varied scale-free and small-world topologies on cellular EA with and without self-adaptation.

Pre-experimental planning. First tests employed the parameter optimization method SPO as recently suggested by Bartz-Beielstein [17]. They revealed that, keeping the population size constant at 400 and the number of connections at 800, in most cases no significant performance increase could be gained by varying the number of offspring per generation or the maximum lifespan of an individual (the latter would lead to a κ-type or comma-type environmental selection/replacement scheme). This also holds for the mutation rate meta-parameter τ needed for self-adaptation, which has therefore been fixed at 0.5. Furthermore, the mutation rate default setting $p_m = 1/l$, with l the representation length, could be verified as a good compromise when using a fixed mutation rate for different problems.

A notable exception is the COUNTSAT problem, where self-adaptation together with large birth surplus and comma-type environmental selection performed very well. However, to simplify interpretation of results, we limited experimentation to plus selection, that is, any parent survives as long as it is not outperformed by its offspring. Our tests also showed that choosing a large population size for the panmictic EA is well-founded for the given problem set, at least when striving for high success rates.

For all problems, we determined suitable run lengths in order to measure success rates that approximate the ones for an infinite number of evaluations. The resulting run lengths are given in Table 1. In most cases, the actual average amount of evaluations needed to reach the global optimum is much lower.

When mutation rates are allowed to change, they still must be initialized with meaningful values. Our testing revealed that either starting with $p_m = 1/l$ or $p_m = 0.5$ for all individuals is advantageous, as opposed to initializing p_m uniformly within $]0, 1[$.

Task. The character of our experiment is explorative; we want to find evidence that helps to answer the questions posed in the Introduction, namely situations in which small-world/scale-free topology based CEAs and/or self-adaptation appear advantageous over a standard, panmictic EA.

Table 1: *Problem designs, common (top) and individual (bottom) part. SR stands for success rate, and AES is the average number of evaluations to solution. Each run was stopped at the given maximum number of evaluations as the only termination criterion.*

Initialization	Number of runs	Performance measures
randomized	100	SR/AES

Problem	Instance	Bits	Max. eval	Optimum
MMDP	20 blocks of 6 bits	120	120000	20.0
wP-PEAKS	10 peaks, $w_1 = 1.0$, $w_{[2\ldots10]} = 0.99$	100	200000	1.0
ECC	12 codes à 12 bits, 12 complementary codes	144	400000	0.0674
COUNTSAT	20 bits	20	120000	1.0

Setup. Utilized problem designs, including initialization and termination criterion, are documented in Table 1. The EA variants employed all use bit-flipping mutation with probability p_m and 2-point crossover. Mating selection is done randomly in the neighborhood of each individual, i.e. uniform selection, or among the whole population for the panmictic variant. We set the crossover probability to 1, so that during each generation, every individual produces one offspring. Replacement —or environmental selection— is performed simultaneously (synchronous) for all individuals, taking the better one of the current individual and its offspring each. Population size (400) and number of connections (800) are kept at CEA standard values to allow for comparison with previous studies [18].

Self-adaptation is performed as suggested by Rudolph [19] for discrete variables, differing only in that a mutation event always flips the accordant bit instead of computing its new value from the old one or choosing it randomly from $\{0, 1\}$. We apply it to the mutation probability only, as depicted in Eqn. 5, where τ is a constant meta-parameter and $N(0, 1)$ stands for a standard normally distributed random variable.

$$p'_{mut} = p_{mut} \cdot \exp\left(\tau \cdot N(0, 1)\right) \tag{5}$$

Thus every individual gets a mutation probability that it bequeathes to approximately half of its children by discrete recombination. We follow the standard scheme of evolution strategies by first applying mutation to the mutation rate, then utilizing the acquired mutation rate for mutating the rest of the genome [8].

Summarizing, four EA variants are run on the test problem set: A panmictic EA, a CEA with fixed mutation rate $p_m = 1/l$, and two CEAs with self-adaptive mutation rates, starting with $p_m = 1/l$ and $p_m = 0.5$, respectively. Except for the panmictic EA, different graphs are tested: For the Watts and Strogatz model topologies tried, we vary the rewiring factor β between 0 and 0.2 . Whereas 0 stands for an unmodified ring structure, $\beta > 0.2$ produces networks that rapidly approach random graphs. The scale-free topologies were created for kernel sizes from the minimum 2 to 28, in which almost half of the available connections must be spent for the kernel, so that at least one per remaining node is left for preferential attachment. With the given parameters, actual topologies have been created anew for every single run.

Experimentation/Visualization. Due to space limitations, we only depict SR (success rate) results for the four problems, see Figures 2 and 3. Table 2 additionally provides

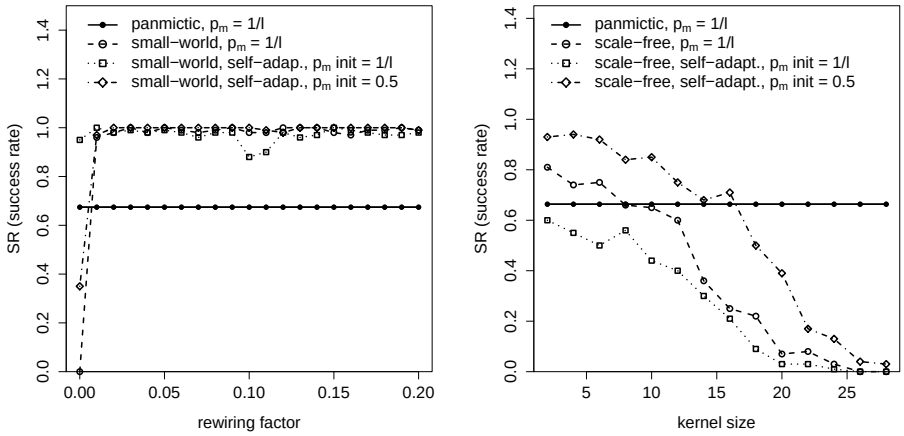


Fig. 2: *Success rates for small-world (left) and scale-free (right) topology CEAs, compared to a similarly parametrized panmictic EA, on the MMDP. Each point is generated from 100 runs.*

numerical values obtained for AES (average number of evaluations to solution) and SR criteria on the MMDP, tables for the other problems are omitted for the same reason.

Observations. The first thing to note is that the SR performance curves look very different for the four test problems. We therefore decided to describe the obtained results separately.

MMDP: Success rates for all small-world topology CEAs (except when $\beta = 0$) are near 1 and thus much higher than 0.66 of the panmictic variant. At the same time, they are a lot slower than the panmictic EA (see Table 2). Ring topology CEAs, i.e. $\beta = 0$, may have failed to succeed because their time consumption would have been even higher than the given limit. Scale-free CEAs with small kernels perform comparable to small-world CEAs with medium rewiring factor, in success rates as well as in speed. For larger kernels, success rates drop dramatically, even below the ones for the panmictic EA. Simultaneously, the length of successful runs increases. Self-adaptation of mutation rates works well in all small-world CEAs and quite good for scale-free CEAs with small kernels. It remarkably lowers the AES if started with $p_m = 1/l$. Interestingly, it was observed that learned mutation rates, especially when started at 0.5, tend to develop towards both ends of the allowed interval, namely 0 and 1, within the same population.

ECC: For both topology types, the fixed mutation CEA outperforms all other variants with respect to the SR criterion. Rewiring rates and kernel sizes seem to have little influence here. The panmictic EA is slightly faster but achieves much worse success rates. Self-adaptation does not seem to work at all for this problem, it delays the CEAs while also reducing success rates.

COUNTSAT: Only one of the four algorithms is able to solve the problem with non-significant success rates: The self-adaptive CEA starting with $p_m = 0.5$. Topology differences seem to have little influence. Unfortunately, we did not try a panmictic

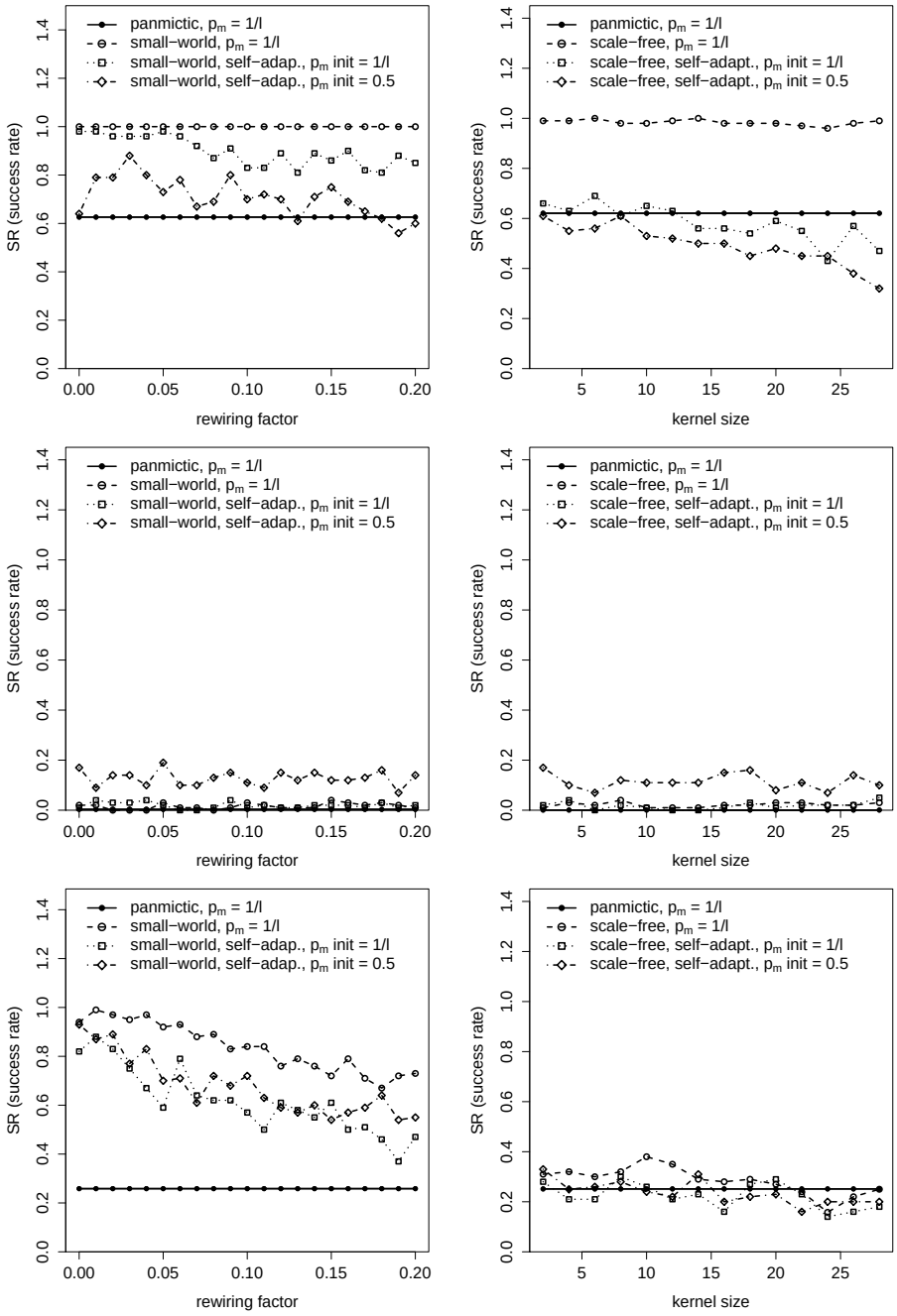


Fig. 3: *Success rates for small-world (left) and scale-free (right) topology CEAs, compared to a similarly parametrized panmictic EA, on problems (top to bottom) ECC, COUNTSAT and wP-PEAKS. Points are generated from 100 runs each.*

Table 2: *Panmictic versus parametrized scale-free and small-world topology cellular EA on problem MMDP. Performance values for EA variants are ordered into blocks of three rows, giving success rates (SR), average evaluations to solution (AES) and AES standard deviations, respectively. Success rates are averaged from 100 runs, AES values and standard deviations computed from the fraction reaching the global optimum. Characteristic path lengths (cpl) and clustering coefficients (C) are determined empirically.*

EA variant: panmictic, non-adaptive, initial $p_m = 1/l$							
SR	0.66	0.66	0.66	0.66	0.66	0.66	0.66
AES	28894	28894	28894	28894	28894	28894	28894
AES std.dev.	11727	11727	11727	11727	11727	11727	11727
kernel sizes ⇒	2	4	6	10	14	20	28
cpl/C ⇒	3.7/0.05	3.7/0.06	3.6/0.07	3.4/0.12	3.3/0.20	3.1/0.39	3.1/0.76
EA variant: scale-free, non-adaptive, initial $p_m = 1/l$							
SR	0.81	0.74	0.75	0.65	0.36	0.07	0.0
AES	58420	59027	57307	65785	65500	74571	—
AES std.dev.	10507	12162	8016	9998	9960	12816	—
EA variant: scale-free, self-adaptive, initial $p_m = 1/l$							
SR	0.60	0.55	0.50	0.44	0.30	0.03	0.0
AES	40067	38473	38480	40364	41067	44000	—
AES std.dev.	10334	2878	3407	3675	3714	4320	—
EA variant: scale-free, self-adaptive, initial $p_m = 0.5$							
SR	0.93	0.94	0.92	0.85	0.68	0.39	0.03
AES	64129	63745	62348	67365	71706	84051	110000
AES std.dev.	10373	10592	9239	11950	14014	15416	10198
rewiring factor ⇒	0.0	0.01	0.02	0.05	0.10	0.15	0.20
cpl/C ⇒	50.4/0.5	15.5/0.47	10.8/0.45	7.4/0.37	5.9/0.28	5.4/0.20	5.0/0.16
EA variant: small-world, non-adaptive, initial $p_m = 1/l$							
SR	0.00	0.96	0.98	1.00	0.98	0.98	0.99
AES	—	100960	84850	70320	61460	56570	55270
AES std.dev.	—	9652	9376	6240	5295	4989	7614
EA variant: small-world, self-adaptive, initial $p_m = 1/l$							
SR	0.95	1.00	0.98	0.99	0.88	1.00	0.98
AES	94880	59220	52420	45030	40200	40200	37320
AES std.dev.	1081	5871	4609	3622	3790	7180	2489
EA variant: small-world, self-adaptive, initial $p_m = 0.5$							
SR	0.35	0.97	1.00	1.00	1.00	1.00	0.99
AES	109540	88920	81040	68680	61380	58480	56040
AES std.dev.	8842	1223	1180	7854	7001	6400	6780

EA with self-adaptation to see if topology has an effect at all. Our impression is that this is not the case but success rather depends on high mutation rates.

wP-PEAKS: Here, the small-world CEAs clearly dominate the panmictic EA, with the fixed mutation rate CEA performing best. Self-adaptation only lowers the success rates. Measured AES values for all small-world variants are largely constant and around 2 to 3-times higher than for the panmictic variant, regardless of the rewiring factor. The scale-free CEAs achieve no better success rates than the panmictic EA, but also require 2 to 3-times more evaluations than the panmictic.

Interpretation. At a first glance, it seems hard to perceive a clear trend within the obtained results. The most we can state is that scale-free topologies do not seem to provide a worthwhile alternative to panmictic or Watts-Strogatz small-worlds. Nevertheless, when thinking about the properties of the utilized test problems and linking them to the algorithm properties of the EA variants regarded as most successful (Ta-

ble 4), we may derive some generalizable conjectures. It seems that problems with a certain degree of separability may profit from localizing operators. However, this also happens for the wP-PEAKS problem which is non-separable.

Table 3: *Test problem properties next to algorithm properties found successful. The ECC problem is not fully separable but organized in blocks. Solution permutable means that fitness of a solution, or subsolution in case of the MMDP, solely depends on the number of ones, not their location.*

Problem	Separable	Solution permutable	Topology matters	Self-adaptation works
MMDP	X	X	X	X
ECC	partly	–	X	–
COUNTSAT	–	X	–	X
wP-PEAKS	–	–	X	–

Concerning self-adaptation, the picture is much clearer. Within our experiments, it worked well for problems with permutable best solutions. That is, several optimal solutions exist that share the number of ones, either in the whole genome as for the COUNTSAT problem, or in the separate building blocks as for the MMDP. Lacking further investigations, we can only speculate why self-adaptation provides an advantage, or at least does not diminish optimization success here. Possibly, the temporary appearance of several different, namely higher mutation probabilities in the course of the optimization process leads to better results.

5 Conclusions

The results of this empirical study indicate that small-world topologies allow for a trade-off between robustness and speed of the search; this is in agreement with the results of [5] on selection pressure, especially when Watts–Strogatz networks are used. In terms of success rate, these population topologies behave at least as well, and often better, than the panmictic case. However, their convergence speed is lower. This effect had already been reported in the case of regular lattice population structures for the same class of problems [18].

On the other hand, scale-free topologies do not seem very helpful in their current form, especially for large kernel sizes. Smaller clique sizes work better but, overall, they do not outperform the standard panmictic setting. This confirms that the selection pressure induced by these topologies on the population may be too high, similar to the panmictic, thus causing premature convergence [5]. However, we have only experimented with static scale-free topologies: we feel that playing with highly connected nodes in a graph would open new perspectives in the control of the exploration/exploitation trade-off, and we intend to try out these ideas in the future.

As far as the EA strategies are concerned, self-adaptation helps if the solution/subsolution is permutable, while fixed mutation performs best overall. In the future we also intend to extend the investigation to continuous problems and to study the dynamics of birth surplus (comma) strategies.

References

1. E. Cantú-Paz. *Efficient and Accurate Parallel Genetic Algorithms*. Kluwer Academic Press, 2000.
2. M. Tomassini. *Spatially Structured Evolutionary Algorithms: Artificial Evolution in Space and Time*. Springer, Berlin, Heidelberg, New York, 2005.
3. R. Albert and A.-L. Barabasi. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.
4. M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.
5. M. Giacobini, M. Tomassini, and A. Tettamanzi. Takeover time curves in random and small-world structured populations. In H.-G. Beyer at al., editor, *Proceedings of the Genetic and Evolutionary Computation Conference GECCO'05*, pages 1333–1340. ACM Press, 2005.
6. M. Preuss and C. Lasarczyk. On the importance of information speed in structured populations. In Xin Yao et al., editor, *Parallel Problem Solving from Nature, PPSN VIII*, pages 91–100. Lecture Notes in Computer Science Vol. 3242, Springer-Verlag, 2004.
7. T. Bäck. Self-adaptation in genetic algorithms. In F. J. Varela and P. Bourgine, editors, *Toward a Practice of Autonomous Systems – Proc. First European Conf. Artificial Life (ECAL'91)*, pages 263–271, Cambridge MA, 1992. The MIT Press.
8. H.-G. Beyer and H.-P. Schwefel. Evolution strategies: A comprehensive introduction. *Natural Computing*, 1(1):3–52, 2002.
9. D. Whitley, S. Rana, J. Dzubera, and K. E. Mathias. Evaluating evolutionary algorithms. *Artif. Intelligence*, 85:245–276, 1997.
10. D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997.
11. D. E. Goldberg, K. Deb, and J. Horn. Massively multimodality, deception and genetic algorithms. In R. Männer and B. Manderick, editors, *Parallel Prob. Solving from Nature II*, pages 37–46. North-Holland, 1992.
12. K. A. De Jong, M. A. Potter, and W. M. Spears. Using problem generators to explore the effects of epistasis. In T. Bäck, editor, *Proceedings of the Seventh ICGA*, pages 338–345. Morgan Kaufmann, 1997.
13. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, 1977.
14. H. Chen, N. S. Flann, and D. W. Watson. Parallel genetic simulated annealing: A massively parallel SIMD algorithm. *IEEE Transactions on Parallel and Distributed Systems*, 9(2):126–136, 1998.
15. S. Droste, T. Jansen, and I. Wegener. A natural and simple function which is hard for all evolutionary algorithms. In *IEEE International Conference on Industrial Electronics, Control, and Instrumentation (IECON 2000)*, pages 2704–2709, Piscataway, NJ, 2000. IEEE Press.
16. D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.
17. T. Bartz-Beielstein. *New Experimentalism Applied to Evolutionary Computation*. PhD thesis, University of Dortmund, April 2005.
18. B. Dorronsoro, E. Alba, M. Giacobini, and M. Tomassini. The influence of grid shape and asynchronicity on cellular evolutionary algorithms. In *2004 Congress on Evolutionary Computation (CEC 2004)*, pages 2152–2158. IEEE Press, Piscataway, NJ, 2004.
19. G. Rudolph. An evolutionary algorithm for integer programming. In Y. Davidor, H.-P. Schwefel, and R. Männer, editors, *Proc. Parallel Problem Solving from Nature – PPSN III, Jerusalem*, pages 139–148, Berlin, 1994. Springer.