Network Optimization
Using Evolutionary Strategies

F. Schweitzer, W. Ebeling, H. Rosé, O. Weiss

Institute of Physics, Humboldt University
Invalidenstr. 110, 10115 Berlin, Germany

e-mail: frank@summa.physik.hu-berlin.de

1 Introduction

The optimization of networks which connect a given set of nodes is of common interest in many different areas, among them electrical engineering (e.g. for electronic platines), telecommunication, or road construction and trade-logistics.

Usually, the optimization has to consider different demands, e.g.:

(i) to minimize the total costs of establishing the network, which should be proportional to the total length of the links between the different nodes,

(ii) to minimize the time to reach any given node from any other node, which should be proportional to the length of the links along the shortest path between the two nodes.

Considering only the first demand, the solution is given by a network where every node is connected to the net by just one (the possible shortest) link, which leads to the *minimal link system*, also known as *minimal spanning tree*. Contrary, considering only the second demand, the solution is given by a network where every node is connected to every other node by a direct link, which leads to the *direct link system* (compare also Fig. 1).

Both optimized solutions have their disadvantages: in the first case, a minimized total length of the network means a very long connection between arbitrary nodes, which includes large detours, since the connection on the existing network exceeds the metric distance by far. Moreover, a minimal link system might be susceptible to breakdown, because every node has only one connection to the net. In the second case, the metric distance is equal to the length of the link, leading to the shortest possible connection between two points – however the total length of the network reaches a maximum.

For practical applications, both solutions could be sufficient under certain circumstances. A minimal link system is appropriate e.g. if the connection distance can be passed with a very high speed. Then the time to reach a given node does not count, and a detour on the network could be easily accepted. On the other hand, a direct link system will be appropriate if the costs of establishing a network do not count compared to the traveling time which should be as short as possible.

Compared to these two idealized limiting cases, in most real applications a network has to be established which minimizes the total length of the connections *as well* as the traveling time. However, a minimized detour means a direct connection between all nodes and therefore, a maximum total length of the network, and minimal costs for the network mean the smallest possible total length for the links. Since both demands could not be satisfied at the same time, we have to find a compromise between the two cases discussed above.

Optimization problems like this are known as *frustrated problems* (Ebeling *et al.*, 1990, 1994). The frustrated optimization problem is characterized by a tremendous number of nearly evenly matched solutions which have to be found in a very rugged landscape of the related optimization function. In order to find some of these matched solutions, evolutionary algorithms are applied (Rechenberg, 1994, Holland, 1975, Schwefel, 1981, Goldberg, 1989, Fogel, 1995). These algorithms are a special class of stochastic search strategies in an ensemble of searchers which adapt certain features from natural evolution. The examples discussed here, are the Boltzmann- and the Darwin strategy, as well as a mix of both of them (Ebeling, Engel, 1986, Ebeling, 1990, Boseniuk *et al.*, 1987, 1991).

With respect to the optimization of networks, we investigate (i) the evolution of the network and the related fitness function during the optimization process, (ii) different optimized solutions (graphs of varying density) for the network in dependence on the degree of frustration.

2 Evaluation of Networks

In order to optimize networks, we first have to define a potential function (or a *fitness function*) which evaluates a given network.

Let us consider a set of nodes which shall be connected by straight lines representing the links. The number of possible graphs g to connect a given set of nodes $p_1...p_N$ is of the order $2^{N(N-1)/2}$. Each graph should be evaluated due to the following potential function:

$$U(g) = (1-\lambda)D(g) + \lambda L(g)\,;\quad 0 \le \lambda \le 1 \tag{1}$$

Here, $D(g)$ represents the mean detour to reach different nodes on the existing network, whereas $C(g)$ represents the total costs for establishing the network, which are related to the total length of the links.

In order to minimize the potential $U(g)$, both terms should be minimized. As discussed above, this means a frustrated problem, since both demands cannot be totally satisfied at the same time. The demand for a minimized detour between every two points represents a *local constraint* to the network, whereas the demand of a minimized total length of the network means a *global constraint*. For example, if we discuss the optimization of a road network (Schweitzer *et al.*, 1995), the local constraints represent the interest of the users who don't like detours, and the global constrains are given by the interests of the government which has to pay for the road construction and therefore tries to minimize it.

The parameter λ is introduced to weight between the two contradicting demands. The case $\lambda \to 0$ leads to a potential function, which only minimizes the detour regardless of the costs of the network, and finally results in the *direct link system* (Fig. 1a). In the opposite case, $\lambda \to 1$, only the costs, which should be proportional to the length of the network will be minimized, which finally leads to the *minimal link system* (Fig 1b). (We note, that in the node system considered here, the minimal link system will be different from the known Steiner tree (Fig.1c), since we assume that any link between nodes should be a straight line and no additional sub-nodes should be constructed to connect given nodes.) If λ is different from 0 or 1, it should be a measure for the *degree of frustration* of the problem.

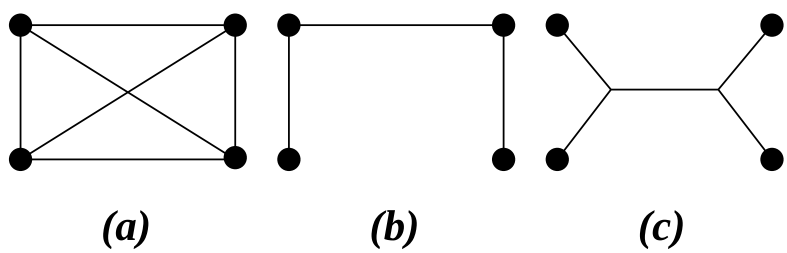

Figure 1: (a) direct link system, (b) minimal link system, (c) Steiner tree for a set of four nodes

In this paper, the direct link system will be used as a reference state, indicated by the symbol $\star$. In this case a direct connection between any two nodes exists. Then, the detour $D(g^\star)$ is zero and the total length of the network gets its maximum value $L^\star$:

$$L^\star = \frac{1}{2}\sum_{i,j=1}^{N} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} = \frac{1}{2}\sum_{i,j=1}^{N} l^\star_{i,j} \quad ; \; D(g^\star) = 0 \tag{2}$$

x_i, y_i are the coordinates of node i in a two-dimensional plane. For values of $0 \leq \lambda \leq 1$, the mean detour $D(g)$ is defined as follows:

$$D(g) = \frac{1}{2}\sum_{i,j=1}^{N} h_{i,j} - l^\star_{i,j} \tag{3}$$

where $l^\star_{i,j}$ means the direct (metric) distance between the nodes i and j (eq. 2) and $h_{i,j}$ is the length of the *shortest route* which connects nodes i and j *on the existing graph.* Obviously, $h_{i,j} = l^\star_{i,j}$ yields for a graph representing the direct link system.

The expression $D(g)$ can be normalized by the total length of the direct link system, which is a known constant for a given set of nodes,

$$d(g) = \frac{D(g)}{L^\star} = \frac{1}{2} \sum_{i,j=1}^{N} \frac{h_{i,j} - l^\star_{i,j}}{L^\star} \tag{4}$$

In order to specify the term representing the costs, we assume the costs simply proportional to the total length of the graph

$$C(g) = \frac{1}{2} \sum_{i,j=1}^{N} l_{i,j} \tag{5}$$

where $l_{i,j}$ is the length of the direct connection between nodes i and j, *if there is any on the existing graph.*

After a normalization of $C(g)$ similar to eq. (4), the potential function describing the network reads finally

$$u(g) = \frac{U(g)}{2L^\star} = \frac{1}{2L^\star} \sum_{i,j=1}^{N} (1-\lambda)\,(h_{i,j} - l^\star_{i,j}) + \lambda\, l_{ij} \tag{6}$$

For the reference state (direct link system), the values are $d(g^\star) = d_{min} = 0$ for the mean detour, $c(g^\star) = c_{max} = 1$ for the costs, and $u(g^\star) = \lambda$ for the potential. Based on eq. (6), a hypothetic network will be optimized with respect to the mean detour and the total length of the network.

3 Boltzmann and Darwin Strategies

An optimization process can be described as a special search for minima of a related potential in the configuration space. In the considered case, the configuration space is defined by the number of all possible graphs g to connect the given set of nodes $p_1 ... p_N$, and the potential is given by $u(g)$ (eq. 6).

Let us consider a numbered set of states $i = 1, ..., s$ in the configuration space, each of them characterized by a scalar U_i (the potential energy). Further, we assume a total number of N searchers participating in the search for potential minima. Then, $N_i(t)$ gives the actual number of searchers occupying the state i at time t. This occupation number should be an integer, but in the limit $N \to \infty$ the occupation fraction $N_i(t)/N$ may be replaced by a probability of occupation at time t denoted by $p_i(t)$.

The optimization process has to ensure that the occupation probability for the minima of the potential will increase during the search. However, since due to the frustration of the problem a large quantity of suitable minima exist, the search strategy should also avoid a total locking of the searchers in the minima found so far in order to guarantee further search.

There are different optimization routines, like the METROPOLIS algorithm (METROPOLIS *et al.*, 1953) or the *simulated annealing* approach (ANDRESEN, 1989, NULTON, SALAMON, 1988) which fulfill these criteria. A dynamics which finds the minimum among the set of scalars (the minimum potential energy) is given by

$$\frac{dp_i(t)}{dt} = \sum_{i \neq j} A_{ij}\, p_j(t) - A_{ji}\, p_i(t) \tag{7}$$

A_{ij} denotes the transition probability for the searcher to move from state i to state j and is defined as follows:

$$A_{ij} = A^0_{ij} * \begin{cases} 1 & \text{if } U_i < U_j \\ \exp\left(-(U_i - U_j)/T(t)\right) & \text{if } U_i \geq U_j \end{cases} \tag{8}$$

This means that transitions towards a deeper minimum in U are always accepted, but transitions which lead to a deterioration are accepted only with a probability related to the difference in the potential. Thus, due to the motion along the gradients the steepest local descent of the potential will be reached; however, due to thermal fluctuations locking in those local minima will be avoided.

The prefactor A^0_{ij} is symmetrical ($A^0_{ij} = A^0_{ji}$), it defines a set of possible states j which can be reached from state i. The simplest definition might be

$$A^0_{ij} = \begin{cases} 1 & \text{if} \quad i \text{ is adjacent to } j \\ 0 & \text{if} \quad i \text{ is not adjacent to } j \end{cases} \tag{9}$$

The term *adjacent* means here that state j results only from a single elementary mutation of state i, in other words, a change between the different states can only occur in small steps.

Similar to *simulated annealing*, the temperature $T(t)$ decreases during the search by a certain rule, e.g. by a power law. This decrease leads to the consequence that first the larger basins of the potential minima are explored ("coarse grained search") and later on a "fine grained" search occurs within these minimum regions.

For constant temperatures, the stationary solution of eq. 7, $p^0_i = \lim_{t \to \infty} p_i(t)$, is known to be the canonical or BOLTZMANN distribution

$$p^0_i \sim \exp\left(-U_i/T\right) \tag{10}$$

therefore we call this optimization strategy BOLTZMANN strategy. It has occured during the cosmic evolution to optimize certain thermodynamic functions. Since the minimum of the potential has the hightest probability, the BOLTZMANN process asymptotically finds the minimum in a given set of scalars U_i. One can show (FEISTEL, EBELING, 1989), that during the search process the function

$$K(t) = \sum_i p_i(t) \log \frac{p_i(t)}{p^0_i} = \frac{F(t) - F_0}{T} \tag{11}$$

is monotonically decreasing. Here $F(t)$ has the meaning of a *free energy* of the system with the equilibrium value F_0.

During biological evolution, some new elements occured in the evolutionary optimization process, namely (i) mutation processes due to error reproduction, and (ii) self-reproduction of species with a fitness above the average fitness, which increases its precision in the course of evolution. We adapt these elements for our search strategy, which is then called DARWIN strategy because it includes some biological elements known from population dynamics (EBELING, FEISTEL, 1982, EBELING *et al.*, 1990, 1994, BOSENIUK *et al.*, 1987, 1991).

Let us consider again the population of N searchers, which are now distributed in different subpopulations $x_i = N_i/N$; $(i = 1, ..., N)$, each characterized by a replication rate E_i which might be proportional to the fitness. Then the average replication rate $\langle E \rangle$ is given by

$$\langle E \rangle = \frac{1}{N} \sum_{i=1}^{N} E_i x_i(t), \qquad N = \sum_{i=1}^{N} x_i(t) \tag{12}$$

Due to the EIGEN – FISCHER dynamics, the evolution of the subpopulations is given by the equation:

$$\frac{dx_i}{dt} = (E_i - \langle E \rangle)\, x_i + \sum_{j \neq i} [A^0_{ij}\, x_j - A^0_{ji}\, x_i] \tag{13}$$

Here the transition rates A_{ij} are assumed to be symmetric, since there are no directed mutations. The effect of an increasing precision in self-reproduction can be considered again by a temperature dependence of the transition rates, where a decreasing temperature leads to a smaller probability of mutation. For $A^0_{ij} \rightarrow 0$, this evolutionary dynamics is known to approach asymptotically a final state where the average fitness $\langle E \rangle$ is equal to the maximal fitness, which means that only the (one) subpopulation with the best fitness value will survive. For finite mutation rates, $A^0_{ij} > 0$, the target of the search is the eigenvector of eq. (13) corresponding to the highest eigenvalue, which for small mutations rates is close to the maximal value E_{max}.

To compare both strategies (EBELING, ENGEL, 1986, BOSENIUK, EBELING, 1991) we note that the BOLTZMANN strategy is able to detect the appropiate potential minima even in a unknown, rugged landscape as long as the potential barriers between local minima are not too high, which forces the locking in side minima. On the other hand, the DARWIN strategy is able to cross high barriers by tunneling if the next minimum is close enough.

In order to combine the advantages of both stragegies, a *mixed* BOLTZMANN–DARWIN *strategy* has been introduced (EBELING, ENGEL, 1986) Here, the asymmmetric transition probabilities (eq. 8) are adopted which favor the transition towards the minimum. On the other hand, the fitness E_i of the subspecies i is chosen to be the negative of the potential U_i indicating that the subspecies which has found the better minimum in the potential landscape, also has the higher reprodution rate. Due to the comparison with the mean value $\langle U \rangle$, there exist a global coupling between the different subpopulations.

It has been shown recently (BOSENIUK, EBELING, 1991, ASSELMEYER, EBELING, 1996) that a *mixed* BOLTZMANN–DARWIN *strategy* will be more successful in solving frustrated problems than both of the strategies. The basic equation for the mixed strategy which is appropriate to solve the optimization problem of minimizing U_i reads

$$\frac{dx_i}{dt} = \kappa(\langle U \rangle - U_i)\, x_i + \sum_{j \neq i} [A_{ij}\, x_j - A_{ji}\, x_i] \tag{14}$$

with the transition matrices A_{ij} obtained from eq. (8). By changing the parameters κ and T in the range

$$0 \leq \kappa \leq 1, \quad 0 < T \leq \infty \tag{15}$$

we may interpolate between the two limit cases

$$\begin{array}{lll} \kappa = 0, & T > 0 & \text{BOLTZMANN strategy} \\ \kappa = 1, & T \to \infty & \text{DARWIN strategy} \end{array}$$

In the discussions above, the discrete numbers of searchers, $N_i(t)$, have been replaced by occupation probabilities $p_i(t)$ or population densities $x_i(t)$. We note that for small numbers N_i the stochastic search process could be also reformulated in terms of a master equation (FEISTEL, EBELING, 1989; SCHWEITZER *et al.*, 1995).

4 Results of Computer Simulations

In order to show the evolution of the network during the optimization process (Fig. 2), we start the computer simulations with an initial graph of 39 nodes close to a direct link system (λ =0.975). For the initial state, the mean detour is $d(g^\star) = 0$, the cost value is $c(g^\star) = 1$, and the potential value is $u(g^\star) = \lambda$. During every time step, the graph is first mutated by adding or removing one link between points and then evaluated. For the optimization, in Fig. 2 the BOLTZMANN strategy is used.

The optimization process occurs in two stages (Fig. 3). Starting with a direct link system, during the first stage the network is strictly thined out. However, as Fig. 3a shows, this does not mean a considerable decrease in the potential, unless a remarkable increase of the mean detour is reached, related to a decrease of the costs. During the second stage, the links between the different nodes are balanced with respect to the costs and the mean detour, resulting in a slowly decrease of both detour and costs. In Fig. 3b, the transition between both stages is marked by the maximum region of the curve.

Fig. 4 presents results for the optimized network in dependence on the frustration parameter λ, which influences the final density of the graph as discussed above.

In Fig. 5 the potential values for the optimized network, u^{opt}, obtained asymptotically are plotted vs. the frustration parameter λ. Surprisingly, the potential minimum in the asymptotic regime is a 4th order power function of λ:

$$u^{opt}(\lambda) = \lambda\{-0.0075\lambda^3 + 0.0144\lambda^2 - 0.0111\lambda + 0.0046\} \tag{16}$$

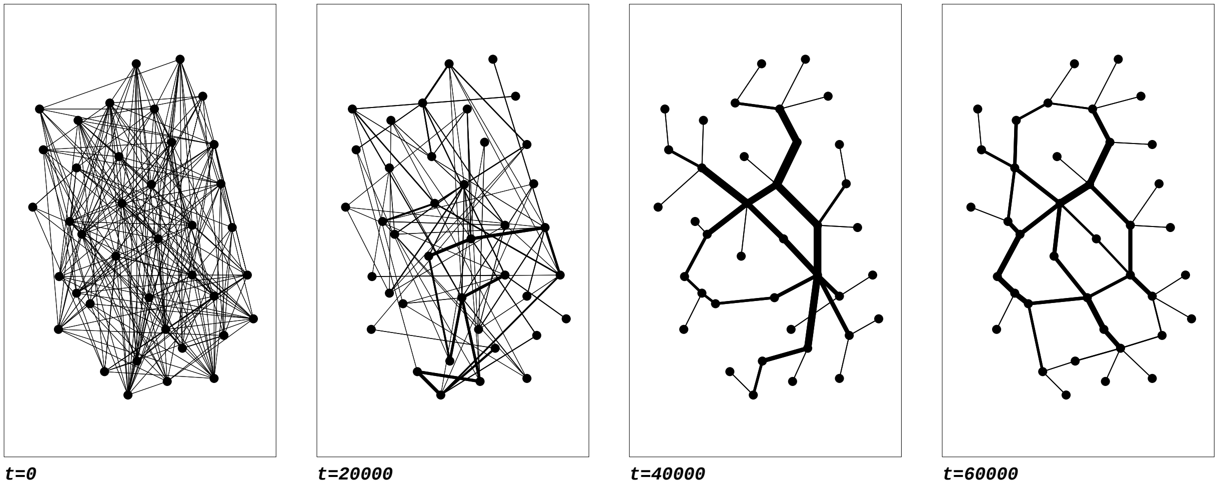

Figure 2: Optimization of a network of 39 nodes ($\lambda = 0.975$) The graph is shown after different time steps. The thickness of the lines indicates how much a given link is used for the shortest possible connection of any node to any other.

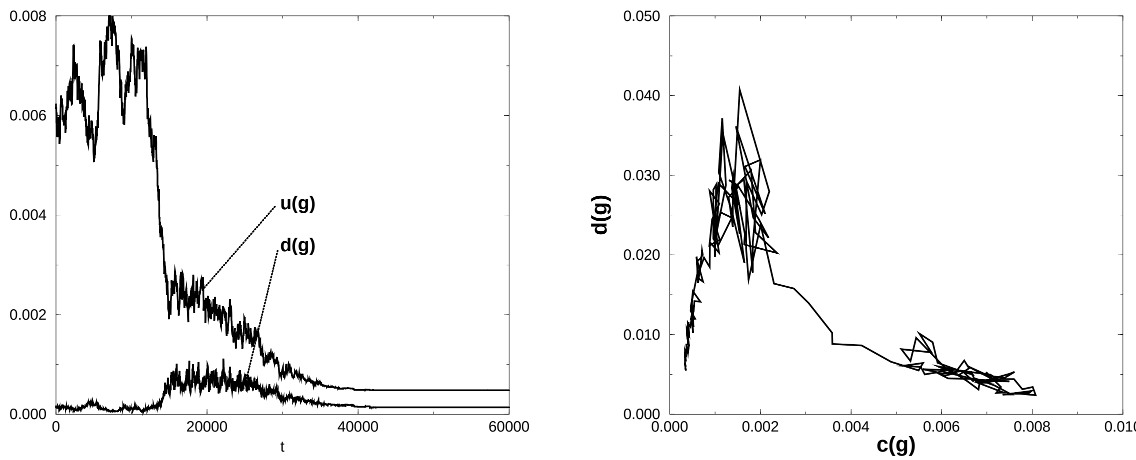

Figure 3: (a) Time dependence of the potenital and the mean detour, (b) mean detour versus costs during the optimization of the network. With respect to the time, the left curve starts in the lower right corner and ends in the lower left corner ($\lambda = 0.975$).

which is also drawn in Fig. 5. This indicates a fixed relation between the asymptotic values of $d(g)$ and $c(g)$ which allows a prediction of the best possible cost and the best affordable detour of the network in dependence on λ.

Finally, we would like to compare the results of the Boltzmann strategy and the mixed strategy which also includes Darwinian elements. As shown in the simulations above, the Boltzmann strategy finds suitable results in the asymptotic limit (about 60.000 simulation steps). However, the mixed Boltzmann-Darwin strategy already finds optimal graphs in a much shorter simulation time, as shown in Fig. 6 for 10.000 simulation steps (obtained for the same number of searchers in both simulations). The optimization function relaxes very fast compared to the Boltzmann curve. With respect to the networks

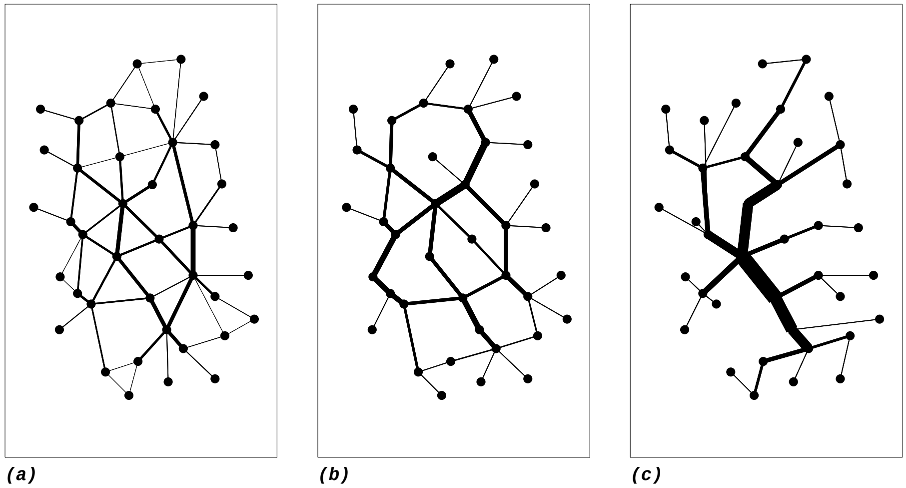

Figure 4: Optimized networks after t=60000 simulation steps for different values of λ: (a)$\lambda = 0.900$, (b)$\lambda = 0.975$, (c)$\lambda = 0.990$

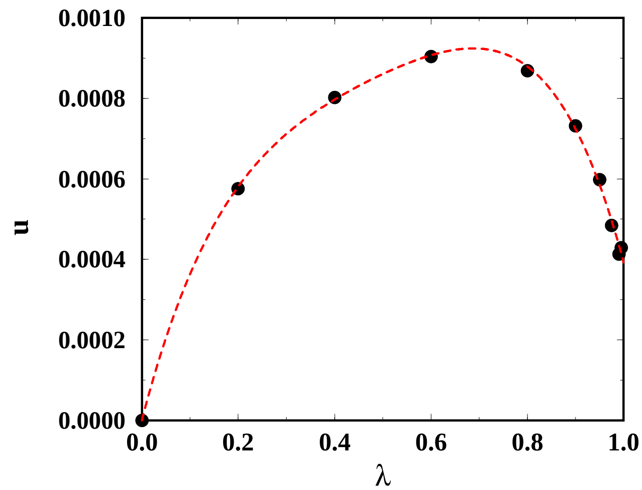

Figure 5: Dependence of the asymptotic potential minimum on the frustration parameter λ. Dots mark the results of optimization simulations, the dashed line is given by eq. (16)

obtained after 10.000 time steps, we find already balanced graphs with the mixed optimization strategy, whereas the graphs obtained from the Boltzmann strategy clearly display failures in optimization.

5 Conclusions

In order to summarize the results presented we want to conclude:

(i) network optimization which has to consider both the connection distance (detour) between different nodes and the total length (costs) of the network, belongs to the

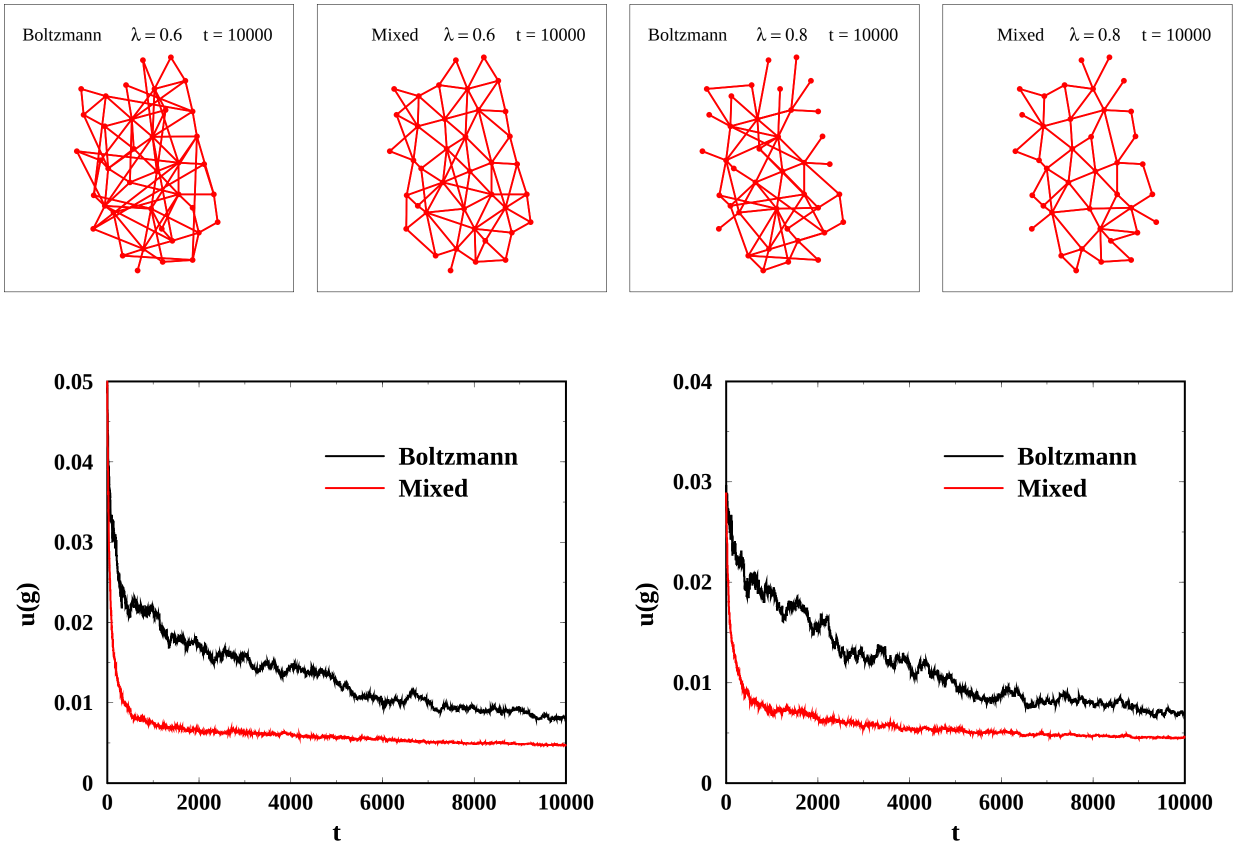

Figure 6: Comparison of Boltzmann and mixed strategies for network optimization. (left) $\lambda = 0.6$ (right) $\lambda = 0.8$ The networks presented are obtained after 10.000 simulation steps, the related potential is displayed below. The ensemble consists of 16 searchers for both strategies.

class of frustrated optimization problems, where numerous evenly matched solutions exist

(ii) evolutionary optimization strategies which include both thermodynamic and biological elements (mixed strategies of simulated annealing, ensemble search, mutation, selection and recombination) provide a suitable tool for finding optimized solutions in relatively short time (preferable in comparison to Boltzmann–like strategies)

(iii) the optimization of networks occurs in two different time scales: (a) thin-out of the network (short time scale), (b) balancing of detour compared to costs (long time scale)

(iv) in the asymptotic limit the potential (fitness) of the optimized network can be described by a power function, which defines a fixed relation between the mean connection distance (detour) and the total lenght (costs) of the network

Acknowledgments

This work has been supported by the Deutsche Forschungsgemeinschaft via SFB 230 "Natural Constructions" (Stuttgart, Germany).

References

B. Andresen: Finite-Time Thermodynamics and Simulated Annealing. Proc. 4th Intern. Conf. Irreversible Processes and Selforganization, Rostock, 1989.

T. Asselmeyer, W. Ebeling: Unified Description of Evolutionary Strategies over continous Parameter spaces. submitted to BioSystems. 1996.

T. Boseniuk, W. Ebeling, A. Engel: Boltzmann and Darwin Strategies in Complex Optimization, Physics Letters 125, 307-310, 1987.

T. Boseniuk, W. Ebeling: Boltzmann-, Darwin-, and Häckel-Strategies in Optimization Problems, Parallel Problem Solving from Nature (editors H.-P. Schwefel, R. Männer), Berlin: Springer, 1991.

W. Ebeling. Applications of Evolutionary Stragegies.*Syst. Anal. Model. Simul.* **7**, 3–16, 1990.

W. Ebeling. A. Engel. Models of Evolutionary Systems and Their Application to Optimization Problems.*Syst. Anal. Model. Simul.* **3**, 377-385, 1986.

W. Ebeling, A. Engel, R. Feistel: Physik der Evolutionsprozesse, Akademie-Verlag, Berlin, 1990.

W. Ebeling, R. Feistel: Models of Darwin Proceses and Evolutionary Principles, BioSystems 15 (1982) 291.

W. Ebeling. H. Rosé. J. Schuchhardt. Evolutionary Strategies for Solving Frustrated Problems, Proc. 1st IEEE Conf. Evolutionary Computation, WCCI Orlando, pp. 79-81, 1994.

R. Feistel, W. Ebeling: Evolution of Complex Systems, Kluwer, Academic Publ, Dordrecht, 1989.

D.B. Fogel: Evolutionary computation – toward a new philosophy of machine intelligence. IEEE Press, (Piscataway NJ), 1995.

D.E. Goldberg: Genetic Algorithms in search, optimization and machine learning, (Readin, MA: Addison-Wesley), 1989.

J.H. Holland: Adaptation in Natural and Artificial Sytstems, University of Michigan Press, Ann Arbor, 1975.

N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller: *J. Chem. Phys.* **21**, 1087, 1953.

J. D. Nulton, P. Salamon: Statistical mechanics of combinatorical optimization. Phys. Rev. **A 37**, 1351, 1988.

I. Rechenberg: Evolutionsstrategie '94, Stuttgart: Frommann-Holzboog, 1994.

F. Schweitzer, W. Ebeling, H. Rosé, O. Weiss, Optimization of Road Networks Using Evolutionary Strategies (Preprint, to be published), 1995.

H.-P. Schwefel: Numerical Optimization of Computer Models, Wiley, New York, 1981.