Hybrid Approach for TSP Based on Neural Networks and Ant Colony Optimization

Dr. Carsten Mueller
University of Economics
Faculty of Informatics and Statistics
Department of Information Technologies
W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic
Email: carsten.mueller@itg-research.net

Niklas Kiehne
ITG Research
Department for Software Engineering
Research Lab for Metaheuristics
Bulmannstreet 48, 90459 Nuremberg, Germany
Email: niklas.kiehne@itg-research.net

***Abstract*—This research article presents a hybrid approach based on an intelligent combination of artificial ants and neurons. Research on different parameter combinations is performed, in order to find the best performing settings. The obtained insights are then subsumed into an intelligent architecture consisting of ACO and SOM.**

I. Introduction

A. Neural Networks and Self Organizing Maps

In biological neural networks the observed topology of neurons is often planar, whereas the input is of multiple dimensions [1], [2]. These neuron topologies do not map the exact input, but rather map the phase space of it. In result, close neurons process those stimuli which are similar. This behaviour is made use of in a special type of neural networks, called Self Organizing Map (SOM) [3], [4].

The application of the biological inspiration is to have a layer of n interconnected artificial neurons that represent the map. Each neuron is associated with a weight vector w_i of the same dimensions as the expected input and a location l_i, typically in the Euclidean plane. The location is used to model the topology of the net. The weights are initialized to either random values or close representations of the expected inputs. The goal is to train the net to respond to similar input vectors within the same region of neurons of the plane. For this purpose, a set of input stimuli M is needed, which is applied successively to the neurons. During training step t, for each $m_i \in M$ the neuron n_s^t with the closest Euclidean distance between stimulus and weight is selected, and called the excitation centre.

Moreover, a set of neurons that are within a range σ^t around the centre are chosen to adjust their weights to the stimulus according to the following formula:

$$w_i^{t+1} = w_i^t + \phi * e^{\frac{-d(l_s,l_i)^2}{2*\sigma^t}} * (m_i - w_i^t) \tag{1}$$

where ϕ is interpretable as the learning rate.

The training consists of epochs in which every $m_i \in M$ is applied exactly once, but in a random order. With each stimulus presentation the time dependent σ^t is updated as $\sigma^{t+1} = \sigma^t * momentum$. The momentum is an adjustable parameter to control how fast the neighbourhood radius decreases over time. After a specified amount of epochs the training is complete.

B. Ant Colony Optimization

Ant Colony Optimization (ACO) is a metaheuristic approach for solving hard combinatorial optimization problems [5], [6]. One important behaviour pattern of ants for ACO is stigmergy, the indirect communication by manipulating the environment [7], [8]. Pheromone trails in ACO serve as distributed, numerical information which the ants use to probabilistically construct solutions to the problem being solved and which the ants adapt during the algorithms execution to reflect their search experience [9], [10], [11].

The behaviour is determined by the parameters α and β [4] - $\alpha > \beta$: there is bigger influence on the choice of path, which is more often explored; $\alpha < \beta$: there is bigger influence on the choice of path, which offers better solution; $\alpha = \beta$: there is balanced dependency between quality of the path and degree of its exploration; $\alpha = 0$: there is a heuristics based only on the quality of passage between consecutive points.

II. Architecture of the Hybrid Approach

The general idea of combining ACO and ANN is to let the ants construct a tour which is then improved by applying a Self Organizing Map. As the ACO algorithm is faster in converging towards a good, but not a very good, solution, the thought is to use the ANN as a kind of local search.

The procedure is as follows:

1) Initialize ACO and SOM with the given parameters
2) Solve the given TSP with the initialized ACO
3) Extract the best found tour in ACO and insert it into the SOM
4) Solve the SOM
5) Return the solution when SOM training is finished

At first, both ACO and SOM are set up with the user specified parameters and the selected TSP case. Then, ACO is started which rapidly scans the search space and finds a useful solution. The solution provided by AS is extracted as a list of cities that depicts the found tour.

Subsequently, the list is handed over to the ANN, that spreads the neuron's weights evenly along the solution. So in direct opposition to the circular layout used in standalone SOM, the weights are distributed across the tour found by ACO.

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.203

At this point, a critical review of the SOM's parameter σ which represents the neighbourhood radius is needed. Once the neurons are dispersed on a valid tour, the usual values, like $\sigma = 3$, would render the inserted solution useless, since the whole structure would be severely deformed. Therefore, a hybrid specific parameter is introduced as the start iteration of the SOM.

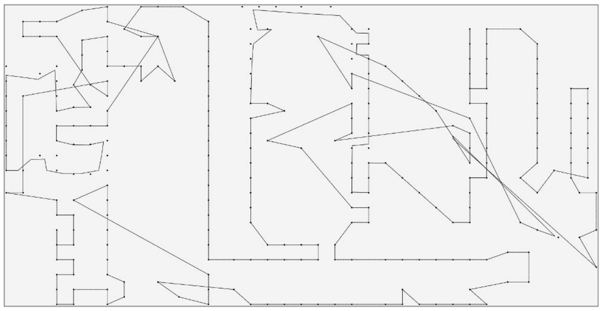

Fig. 1. Second SOM iteration. With two exceptions, the ACO tour is visible, only two regions differ from it

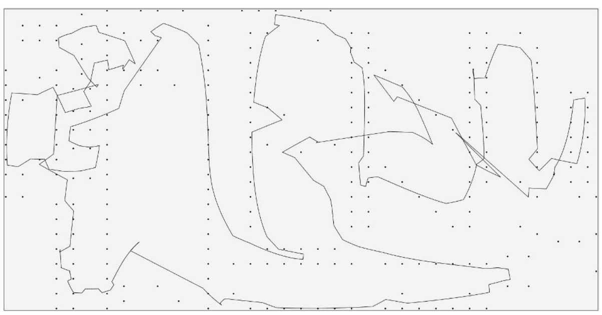

Fig. 2. After 100 iterations the general character of the ACO tour is still observable, but nearly all weights were moved from their original position

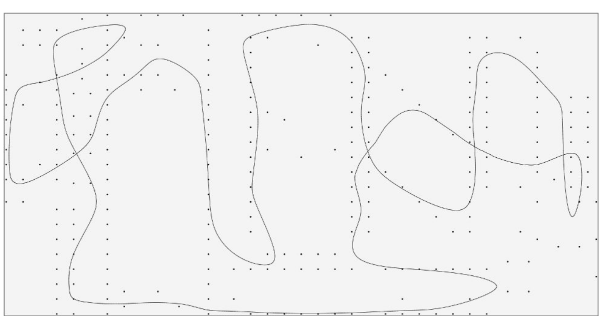

Fig. 3. At iteration 1000 the neurons form a smooth and vague representation of the handed over tour

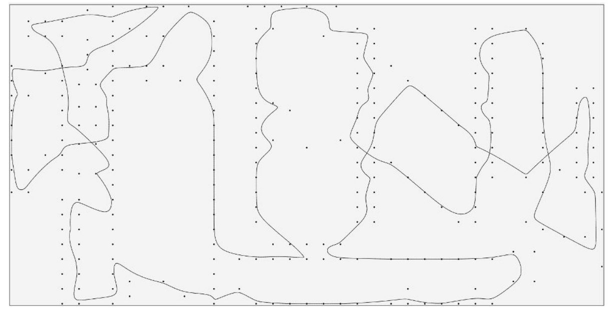

Fig. 4. With iteration 3000 the algorithm starts to move the weights back to the cities

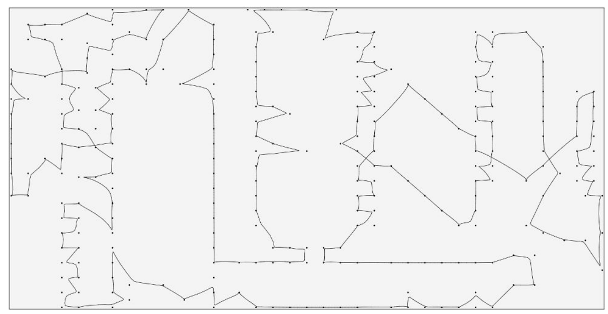

Fig. 5. The algorithm is nearly finished in the 5000th iteration, only few cities are not directly connected

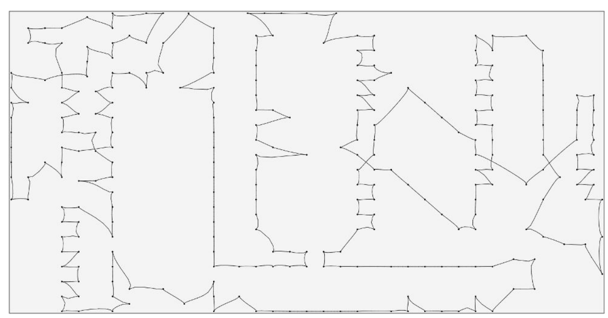

Fig. 6. After 7000 training steps the handed over tour is successfully refined by an ANN. The tour found by ACO had a relative error of 22.9%, whereas the tour after applying SOM achieves 8.0%

The idea is to simulate an advanced progress in the neural net, in which, due to the momentum parameter, changes are only applied to smaller groups of weights.

This is done by computing σ_x with x being the start iteration:

$$\sigma_x = \sigma_0 \cdot momentum^x \quad (2)$$

Different settings of the start iteration effectively change the amount of weights that are slightly detached from the inserted tour, as shown in Figure 2.

The SOM algorithm is then proceeding without further adjustments. As shown in the Figures 1 to 6, the net first relocates the neurons of their initial positions, but only to a limited extent. One effect is, that most of the intersections are straightened out, which is due to the simultaneous movement of multiple weights. In this spirit, the initial softening is some kind of local optimization, whose degree of locality is controllable through a parameter. Once a specific neighbourhood radius is reached, the behaviour of the algorithm is seemingly reversed.

As the amounts of neurons that are moved during one training step is decreasing over time, a behavioural turning point is observed. After the softening reached a peak, the weights are moved back to the nodes, but now without the errors the ants made (see transition from Figure 3 to 4).

The algorithm terminates after the weights are successfully distributed on the cities, which is the known procedure as seen in standalone SOM. In addition, a tournament selection is used, so that the best of both available tours is returned.

In summary, the hybrid approach consists of the sequential processing of a TSP by ACO and SOM. The improved performance emerges from the special application of the neural net, which leads to the local straightening of crossings in the tour supplied by ACO. This behaviour is accomplished through simulating an advanced state in the overall procedure of the SOM algorithm, such that smaller amounts of neurons are moved as compared to the original approach.

The SOM is therefore used as a local optimization technique to refine the tour found by ACO.

III. Parameter Dependencies and Optimization

In the following, the parameters used for fine grained control over the algorithm's performance are explained. Since the proposed architecture uses both ACO and SOM, the parameters available are theoretically the sum of both algorithms.

But examining the influences of all these parameters, and especially their dependencies, is an extensive task. So before the actual evaluation, a logical analysis of the parameters at hand is carried out with the goal of identifying the substantial influences.

At first, parameters concerning the ACO algorithm, such as α and β, are not necessarily parts of the investigation. This is mostly due to the fact, that ACO is used to find the best possible solution in a reasonable time. But since the settings to achieve this behaviour were already figured out in related research, a repeated evaluation can be omitted.

The evaluated standard parameters for the ACO component are $\alpha = 1$, $\beta = 3$, an initial pheromone of 30, 5 ants and 1000 iterations. Furthermore, the examination of the SOM's parameters yielded, that the momentum and the number of neurons factor have good standard values that can be applied through all test cases. Therefore, the chosen values are a momentum of 0.999 and a number of neuron factor of 6. The remaining parameters are σ, ϕ and the number of iterations of the neural net, as well as the hybrid specific start iteration.

- σ: Similar to the functionality and influence of σ in standalone SOM, this parameter represents the neighbourhood radius. But since the ANN is used in a different way than before, good settings are to be discussed. As the neural net is simulated to start in a later iteration, the hybrid approach changes σ depending on the stated start iteration. Figure 9 shows, that best values for σ are between 2 and 8, with different starting iterations respectively. Considering Figure 8 scales the values down to $2 < \sigma \leq 5$.
- ϕ: The meaning of ϕ does not change, but it is still an influential parameter. As seen in Figures 7 and 10, valuable settings are $0.2 < \phi < 0.5$.
- Number of iterations: The used number of iterations which represents the number of times a training stimulus is presented to the neural net has probably the most obvious ranges. All related graphs (see Figures 7-12) state, that the number of iterations should be not less than 2000, with the hint that even higher numbers could result in better solutions.
- Start iteration SOM: The start iteration parameter is controlling how far the neural net is delayed, and therefore closely connected to the behaviour of the hybrid approach. It is responsible for the degree of locality of the local search character of the SOM, since it influences how far neurons are detached from the ACO tour. Choosing a start iteration of zero and a high value of σ would move the weights to such an extent, that the information provided by the ants is lost. But a value too high would cause no improvement at all, since only single weights are adapted per iteration.
- Hence the closeness of minima and maxima as seen in Figure 9. Interestingly, the area in the upper half shows the results of the AS tour, caused by the mentioned single weight adjustment. So the parameter has to be chosen carefully, since even small changes of a few hundred iterations or less may decide on best or worst performance. In general, the start iteration should not be greater 2000 and not lower than 750.

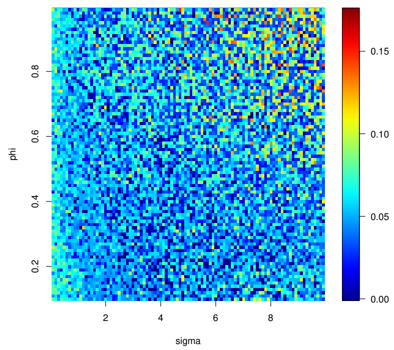

Fig. 7. Parameter test of σ in a range from 0.1 to 10 with a step size of 0.1 and ϕ ranging from 0.1 to 1 with a step size of 0.01

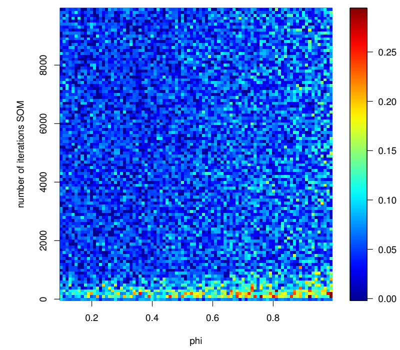

Fig. 10. Parameter test of ϕ in a range from 0.1 to 1 with a step size of 0.01 and the number of iterations in a range from 0 to 10000 with a step size of 100

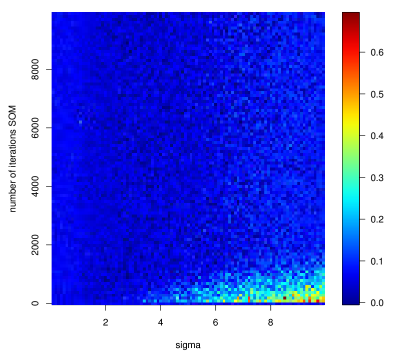

Fig. 8. Parameter test of σ in a range from 0.1 to 10 with a step size of 0.1 and the number of iterations in a range from 0 to 10000 with a step size of 100

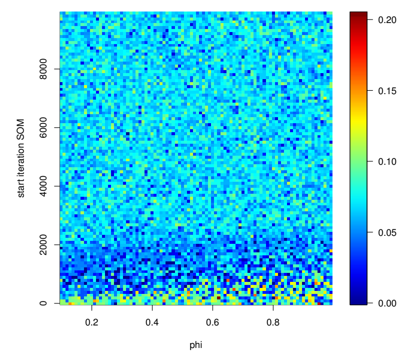

Fig. 11. Parameter test of ϕ in a range from 0.1 to 1 with a step size of 0.01 and the start iteration in a range from 0 to 10000 with a step size of 100

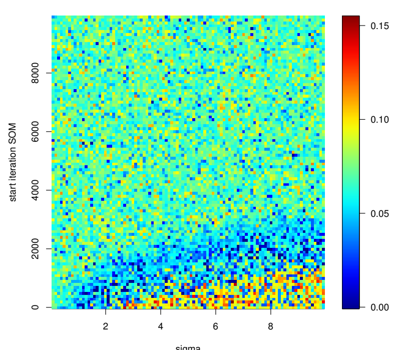

Fig. 9. Parameter test of σ in a range from 0.1 to 10 with a step size of 0.1 and the start iteration of the SOM in a range from 0 to 10000 with a step size of 100

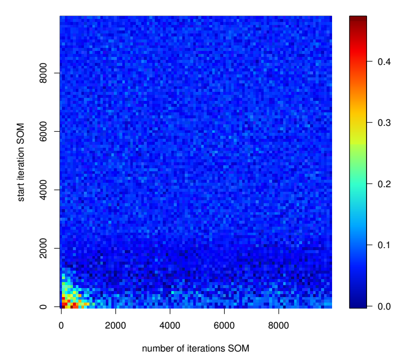

Fig. 12. Parameter test of the number of iterations in a range from 0 to 10000 with a step size of 100 and the SOM's start iteration in a range from 0 to 10000 with a step size of 100

IV. Statistical analysis

In this section the hybrid approach is compared to its components ACO and SOM regarding their performances.

The Welch Two Sample t-test is applied on the acquired results to validate the statistical significance of the improvement. Tables I, II and III show the results of the applied benchmarks as well as the used parameter settings for the most promising performance. Due to the different characteristics of the TSP cases the settings are not constant throughout the instances.

	berlin52	ch130	tsp225	a280	pa561
bestKnownOptimum	7542	6110	3916	2579	2763
bestDistanceFound	7597	6122	3931	2588	2769
numberOfIterations	10000	10000	10000	10000	10000
numberOfAnts	10	10	10	10	10
alpha (α)	1.5	1.6	1.3	1.3	1.5
beta (β)	2.5	2.7	2.5	2.5	2.4
rho (ρ)	0.975	0.980	0.975	0.975	0.99
min	7597	6122	3931	2588	2769
max	8066	6589	4626	3909	4068
mean	7715.65	6224.48	4114.02	2945.36	3133.28
median	7716	6224	4116	2946	3133
sd	68.69	60.18	106.09	207.04	210.67

TABLE I. Statistical Analysis of the performance benchmark for ACO

	berlin52	ch130	tsp225	a280	pa561
bestKnownOptimum	7542	6110	3916	2579	2763
bestDistanceFound	7573	6117	3922	2579	2766
numberOfIterations	10000	10000	10000	10000	10000
sigma (σ)	3.0	2.9	2.8	2.8	3.2
phi (ϕ)	0.33	0.41	0.39	0.39	0.41
momentum	0.999	0.999	0.999	0.999	0.999
min	7573	6117	3922	2579	2766
max	7919	6567	4595	3747	3830
mean	7685.45	6220.18	4100.52	2915.23	3108.59
median	7686	6220	4101	2912	3107
sd	64.29	59.87	103.12	195.25	196.89

TABLE II. Statistical Analysis of the performance benchmark for SOM

	berlin52	ch130	tsp225	a280	pa561
bestKnownOptimum	7542	6110	3916	2579	2763
bestDistanceFound	7544	6110	3916	2579	2765
numberOfIterations	10000	10000	10000	10000	10000
numberOfAnts	10	10	10	10	10
alpha (α)	1.5	1.6	1.3	1.3	1.5
beta (β)	2.5	2.7	2.5	2.5	2.4
rho (ρ)	0.975	0.980	0.975	0.975	0.99
sigma (σ)	3.8	4.2	3.7	3.7	4.1
phi (ϕ)	0.44	0.40	0.43	0.43	0.39
momentum	0.999	0.999	0.999	0.999	0.999
startIteration	1500	1200	1200	1100	1100
min	7544	6110	3916	2579	2765
max	7948	6528	4479	3651	3932
mean	7669.12	6215.62	4093.46	2901.59	3091.92
median	7669	6215	4092	2900	3090
sd	72.36	61.42	103.47	187.39	190.96

TABLE III. Statistical Analysis of the performance benchmark for the Hybrid

	berlin52	ch130	tsp225	a280	pa561
aco and hybrid (p-value)	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *
som and hybrid (p-value)	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *	≤ 0.05 *

TABLE IV. Statistical Analysis - Welch Two Sample t-test

The p-values of the statistical tests allow the statement, that the hybrid approach outperforms ACO and SOM on a level of 5% (as seen in Table IV).

V. Conclusions

In this research article a hybrid approach based on an intelligent combination of artificial ants and neurons was developed. At first, the performance of the hybrid's components was evaluated separately. For this purpose, the respective parameters, their dependencies and influences on the algorithm's behaviour where investigated and explained.

Multiple tests of different parameter combinations were studied, in order to find the best performing parameter settings.

The obtained insights subsumed into an intelligent architecture consisting of ACO and SOM. It was shown, that the proposed hybrid approach outperformed both its components with a high statistical significance.

References

[1] P. Dayan and L. F. Abbott, *Theoretical neuroscience*. MIT Press, 2001.

[2] F. M. Bear and B. W. Connors, *Neuroscience*. Williams and Wilkins, 2007.

[3] S. Camazine, J. L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, *Selforganization in biological systems*. Princeton University Press, 2001.

[4] C. M. Verleysen, "Special Issue on Advances in Self-Organizing Maps," *Neural Networks*, vol. 19(5-6), pp. 721–976, 2006.

[5] M. Dorigo, M. Birattari, and T. Stutzle, "Ant Colony Optimization," *IEEE Computational Intelligence Magazine*, vol. 1, pp. 28–39, 2006.

[6] M. D. Toksari, "Ant colony optimization for finding the global minimum," *Applied Mathematics and Computation*, vol. 176(1), pp. 308–316, 2006.

[7] G. Theraulaz and E. Bonabeau, "A Brief History of Stigmergy," *Artificial Life*, vol. 5(3), pp. 97–116, 1999.

[8] M. Dorigo, E. Bonabeau, and G. Theraulaz, "Ant algorithms and stigmergy," *Future Generation Computer Systems*, vol. 16(9), pp. 851–871, 2000.

[9] M. Dorigo, V. Maniezzo, and A. Colorni, "The Ant System: An autocatalytic optimizing process," Dipartimento di Elettronica, Politecnico di Milano, Italy, Tech. Rep., 1991.

[10] ——, "Positive feedback as a search strategy," Dipartimento di Elettronica, Politecnico di Milano, Italy, Tech. Rep., 1991.

[11] ——, "The Ant System: Optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and CyberneticsPartB*, vol. 26(1), pp. 29–41, 1996.