Trusted Evolutionary Algorithm

Dudy Lim, Yew-Soon Ong, Yaochu Jin, Bernhard Sendhoff

2006

Preprint:

This is an accepted article published in Congress on Evolutionary Computation. The final authenticated version is available online at: https://doi.org/[DOI not available]

2006 IEEE Congress on Evolutionary Computation
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

Trusted Evolutionary Algorithm

Dudy Lim, Yew-Soon Ong, *Member, IEEE,* Yaochu Jin, *Senior Member, IEEE*, and Bernhard Sendhoff, *Senior Member, IEEE*

Abstract— **In both numerical and stochastic optimization methods, surrogate models are often employed in lieu of the expensive high-fidelity models to enhance search efficiency. In gradient-based numerical methods, the trustworthiness of the surrogate models in predicting the fitness improvement is often addressed using *ad hoc* move limits or a trust region framework (TRF). Inspired by the success of TRF in line search, here we present a Trusted Evolutionary Algorithm (TEA) which is a surrogate-assisted evolutionary algorithm that exhibits the concept of surrogate model trustworthiness in its search. Empirical study on benchmark functions reveals that TEA converges to near-optimum solutions more efficiently than the canonical evolutionary algorithm.**

I. Introduction

Over the last decades, Evolutionary Algorithms (EAs) have gained significant interest in diverse areas, including various complex real-world applications, such as aerodynamic airfoil design [1], rotor blade design [2], scheduling [3], art design [4], and flexible space structure design [5]. Their success and popularity lie in their ease of implementation and the ability to locate close to the global optimum designs even on problems with discontinuous surfaces. As a stochastic multi-point search strategy, EA often engages enormous fitness function evaluations before converging to near optimum solutions. In many complex systems, each fitness evaluation may require the simulation of the high-fidelity analysis codes, such as Finite Element Analysis (FEA), Computational Fluid Dynamics (CFD), etc., that can varies from minutes to hours of supercomputing time. Hence, the use of EAs often becomes computationally prohibitive for this class of problems. To enhance the computational efficiency of standard EAs, it is now becoming standard practice to employ computationally cheap surrogate models in place of original fitness function which is computationally expensive. A variety of techniques for the constructions of surrogate model, often also referred to as surrogate model, meta-models or approximation models, have been used in engineering design optimization. Among these techniques, Polynomial Regression (PR, also known as response surface method) [6-8], Artificial Neural Network (ANN), Radial Basis Function (RBF) [9-10], and Gaussian Process (GP) (also referred to as Kriging or Design and Analysis of Computer Experiments (DACE) models) [11-13] are the most prominent and commonly used techniques.

D. Lim is with the Emerging Research Lab, School of Computer Engineering, Nanyang Technological University, Blk N4, B3b-06, Nanyang Avenue, Singapore 639798 (e-mail: dlim@ntu.edu.sg).

Y. S. Ong is with the Division of Information Systems, School of Computer Engineering, Nanyang Technological University, Blk N4, 02b-39, Nanyang Avenue, Singapore 639798 (phone: +65-6790-6448; fax: +65-6792-6559; e-mail: asysong@ntu.edu.sg).

Y. Jin and B. Sendhoff are with the Honda Research Institute Europe GmbH, Carl-Legien-Strasse 30, 63073 Offenbach/Main, Germany (email: {yaochu.jin,bernhard.sendhoff}@honda-ri.de).

Over the recent years, there has been increasing interests on the development of new EA frameworks that employ a diverse of surrogate models for solving computationally expensive problems under a limited computational budget. Hence, there are now various ways to integrate surrogate models into an evolutionary search. In [11], Ratle proposed a strategy for integrating GA with kriging approximation model and uses a heuristic convergence criterion to decide when the model should be updated. El-Beltagy et. al. [14] extends the work by considering the issue of balancing the concerns of optimization with those of Design of Experiments (DOE). Jin et. al. proposed the coupling of ES with neural network models in [15]. The concept of "generation control" and "individual control" in the evolutionary search was introduced. Further, some empirical criteria for switching between the exact fitness function and approximate models throughout the EA search are provided. Other strategies using GP and the idea of pre-selecting portions of the EA population that undergoes exact fitness evaluations were also considered in [16] and [17]. A recent survey paper that outlines some of the fitness approximation models and data sampling techniques typically used in evolutionary computation can also be found in [18].

In gradient-based methods, such as conjugate-gradient, quadratic programming, and steepest descent methods, the trustworthiness of the surrogate models in predicting fitness improvement or controlling approximation errors is typically addressed using *ad hoc* move limits or a trust region framework. The classical trust region framework (TRF) was proposed for managing quadratic Taylor series model in line search [19] and subsequently extended for generalized approximation models [20]. The success of the TRF lies in its ability to predict fitness improvement in the optimization by adaptively controlling the move limits and further guarantees global convergence under mild assumptions on the accuracy of the surrogate model. In contrast, since EAs make use of probabilistic recombination operators, controlling the step size of design changes (to control the accuracy of approximate fitness predictions) is not as straightforward as in gradient-based optimization algorithms.

To date, there have been some notable works on integrating trust region theory into EA. In [21], Ong et. al. propose a surrogate assisted memetic algorithm for solving optimization problems with computationally expensive fitness function and general constraints, on a limited computational budget. The essential backbone of the framework is an evolutionary algorithm coupled with a trust region managed feasible sequential quadratic programming solver in the spirit of Lamarckian learning. The TRF is used for interleaving use of exact models for the objective and constraint functions with computationally cheap surrogate models during local search. Extensions to enhance search efficiency and approximation accuracy using gradient information and multi-level surrogates were also considered in [22] and [13] recently. Inspired by the success of trust region theory in line search and memetic algorithm, we present a Trusted Evolutionary Algorithm (TEA) for solving optimization problems with computationally expensive fitness functions in this paper. A trusted surrogate-assisted evolutionary algorithm or TEA in short is proposed and investigated. In particular, TEA is designed to maintain good trustworthiness of the surrogate models in predicting fitness improvement or controlling approximation errors throughout the evolutionary search.

The remaining of this paper is organized as follows. Section 2 provides a brief discussion on the classical trust region framework for regulating a surrogate-assisted numerical method. In section 3, we present the Trusted Evolutionary Algorithm (TEA) proposed for solving optimization problems with computationally expensive fitness functions. Section 4 presents our empirical study on two highly multi-modal benchmark functions for both low and high dimensions. Numerical comparisons to the canonical EA and TRF are also reported in the same section. Finally, Section 5 concludes this paper.

II. Trust Region Theory in Numerical Optimization

In this section, we present an overview on the trust region theory [23] for numerical optimization. Without loss of generality, we consider the general bound constrained nonlinear programming problem of the form:

$$\begin{aligned} &\textbf{Minimize}: f(\vec{x}) \\ &\textbf{Subject to}: \vec{x}_l \le \vec{x} \le \vec{x}_u \end{aligned} \tag{1}$$

where $f(\vec{x})$ is a scalar-valued objective function, $\vec{x} \in \Re^n$ is the vector of design variables, while $\vec{x}_l$ and $\vec{x}_u$ are vectors of lower and upper bounds for the design variables.

The classical trust region approach starts from a random initial guess to build a quadratic model using Taylor series approximation given by:

$$\hat{f}\left(\vec{x}_b^{\,k} + \vec{d}\right) = f\left(\vec{x}_b^{\,k}\right) + g_k^T \vec{d} + \frac{1}{2} d^T H_k \vec{d} \tag{2}$$

where $f(\vec{x}_b^{\,k})$ represents the original fitness value of the initial guess, $\vec{x}_b^{\,k}$. $\vec{d}$ is any arbitrary step from $\vec{x}_b^{\,k}$. g_k represents the approximated gradient $\nabla f(\vec{x}_b^{\,k})$ while H_k is the approximated second derivative of the exact fitness function. The trust region approach then proceeds with a line search in the region bounded by $\vec{\Delta}^k$ in equation (3) to locate the locally optimum point $\vec{x}_{opt}^k$ by solving:

$$\begin{aligned} &\text{Minimize: } \hat{f}\left(\vec{x}_b^{\,k} + \vec{d}\right) \\ &\text{Subject to: } \left\|\vec{d}\right\| \le \vec{\Delta}^k \end{aligned} \tag{3}$$

In every iteration, k, the exact value of the fitness function, $f\left(\vec{x}_{opt}^k\right)$, is used to compute the figure of merit, ρ^k:

$$\rho^k = \frac{f\left(\vec{x}_b^k\right) - f\left(\vec{x}_{opt}^k\right)}{\hat{f}\left(\vec{x}_b^k\right) - \hat{f}\left(\vec{x}_{opt}^k\right)} \tag{4}$$

ρ^k provides a ratio of the exact to predicted change in the fitness value at the k^{th} iteration. Using ρ^k, the trust radius, Δ^k, is then updated in the following manner:

$$\begin{aligned} \vec{\Delta}^{k+1} &= c_1 \vec{\Delta}^k, \text{ if } \rho^k \le r_1, \\ &= c_2 \Delta^k, \text{ if } r_1 < \rho^k \le r_2, \\ &= c_3 \Delta^k, \text{ if } \rho^k > r_2 \end{aligned} \tag{5}$$

Typical configurations for c_1, c_2, r_1, r_2 are defined as 0.25, 1.0, 0.25, and 0.75, respectively. Further, c_3 is set to 2 if $\| \vec{x}_{opt}^k - \vec{x}_b^k \|_\infty = \Delta^k$ or unity when $\| \vec{x}_{opt}^k - \vec{x}_b^k \|_\infty < \Delta^k$. The reader is referred to [19] and [20] for the details on how these values are obtained heuristically.

In summary, the decision to contract or expand the trust radius depends on the ability of approximation model in predicting fitness improvements. If excellent search improvement is attained within the current search region, it makes sense for one to be more adventurous by expanding the trust radius. On the other hand, when search improvement is only moderate, the same trust radius is maintained. Otherwise, the region in which the model is considered to be trustworthy is reduced since it is well-defined that approximation accuracy improves at region closer to $\vec{x}_b$ for a quadratic model. The search then proceeds to start from the optimized solution, $\vec{x}_b^{k+1}$, or restarts from the previous solution $\vec{x}_b^k$ depending on ρ^k as follows:

$$\begin{aligned} &\vec{x}_b^{k+1} = \vec{x}_{opt}^k, \text{ if } \rho^k > 0 \\ &\vec{x}_b^{k+1} = \vec{x}_b^k, \text{ otherwise} \end{aligned} \tag{6}$$

In the engineering literatures, there exist a variety of techniques for constructing approximation models that produce more accurate prediction than quadratic model particularly on problems with complex surfaces [24-25]. To accommodate the plethora of modeling techniques, Alexandrov et al. extended the classical trust region theory for general approximation model in [20]. It was also shown in [20] that the global convergence of the trust-region framework is ensured when the zero-order and first-order consistency conditions are imposed at the initial guess, i.e., $\hat{f}(\vec{x}_b^k) = f(\vec{x}_b^k)$ and $\nabla\hat{f}(\vec{x}_b^k) = \nabla f(\vec{x}_b^k)$. The generalized trust region framework (TRF) is outlined in Figure 1.

Begin TRF
- Set initial guess of best point, $\vec{x}_b^0$.
- Set k=0
- **While** termination condition is not met:
  - Build a local approximation model around $\vec{x}_b^0$.
  - Find the optimal solution in the approximation model, by solving the sub-problem:

    Minimize: $\hat{f}\left(\vec{x}_b^k + \vec{d}\right)$

    Subject to: $\left\|\vec{d}\right\| \le \vec{\Delta}^k$
  - Compute the figure of merit, ρ^k:

    $$\rho^k = \frac{f\left(\vec{x}_b^k\right) - f\left(\vec{x}_{opt}^k\right)}{\hat{f}\left(\vec{x}_b^k\right) - \hat{f}\left(\vec{x}_{opt}^k\right)}$$
  - Update $\vec{x}_b^{k+1}$ and $\vec{\Delta}^{k+1}$ according to ρ^k.
  - k=k+1.
- **End** while

End TRF

Fig. 1. Trust Region Framework for generalized approximation models.

III. Trusted Evolutionary Algorithm

In this section, we propose a trust-region inspired evolutionary search for solving optimization problems having computationally expensive fitness functions, which we label as the Trusted Evolutionary Algorithm (TEA).

Here, we would like to highlight the conceptual difference between TEA and other surrogate-assisted EA (SAEA). It is worth noting that existing SAEA focuses generally on approximation quality of the surrogate model used in the evolutionary search. Due to the *curse of dimensionality*, global models are increasingly difficult to construct for problems with large number of variables. As a result, many recent SAEA frameworks [21][26] have opted for local over global models to enhance approximation quality. In contrast, we are more interested in predicting search improvement in the context of optimization as opposed to the quality of the approximation, which is a regarded as a secondary objective.

The pseudo code of the proposed TEA is outlined in Figure 2. For the sake of readability, we present the algorithm in two main phases, i.e., the initialization and search phases.

Begin TEA

Initialization Phase:
- Perform canonical EA for some initial generations until a database of sufficient design vectors/fitness values are archived for surrogate modeling.
- Choose m nearest points to current best point, $\vec{x}_b$.
- Determine initial trust radius, $\vec{\Delta} = \Delta_{1,}\Delta_{2,} ..., \Delta_n$ $\Delta_j = \min\left(x_{dj}^{\max} - x_{bj}, x_{bj} - x_{dj}^{\min}\right)$, j=1,2,…,n. n denotes the dimensionality, $\vec{x}_d$ represents points in the central database.

Search Phase:
- **While** (search termination conditions not met)
  - Create approximation model using m nearest points to $\vec{x}_b$ from the database.
  - Create offspring X= $\{\vec{x}_{1,}\vec{x}_2,...,\vec{x}_\lambda\}$ from parent $\vec{x}_b$ using evolutionary operator, i.e. mutation.
  - Find the optimal solution $\vec{x}^{opt}$ for t search iterations using the approximation model, by solving the sub-problem:

    Minimize: $\hat{f}\left(\vec{x}\right)$

    Subject to: $\vec{x}_b - \vec{\Delta} \le \vec{x} \le \vec{x}_b + \vec{\Delta}$
  - Evaluate $\vec{x}^{opt}$ using exact fitness function and archive it in the central database together with the design vector.
  - Determine the figure of merit, ρ.

    $$\rho = \frac{f\left(\vec{x}_b\right) - f\left(\vec{x}_{opt}\right)}{\hat{f}\left(\vec{x}_b\right) - \hat{f}\left(\vec{x}_{opt}\right)}$$
  - Update trust radius using

    $\vec{\Delta} = 0.25\vec{\Delta}$, if $\rho \le 0.25$

    $= \vec{\Delta}$, if $0.25 < \rho \le 0.75$

    $= \zeta\vec{\Delta}$, if $\rho > 0.75$

    $\zeta = 2$, if $\| \vec{x}_{opt} - \vec{x}_b \|_\infty = \vec{\Delta}$ or

    $\zeta = 1$, if $\| \vec{x}_{opt} - \vec{x}_b \|_\infty < \vec{\Delta}$

    and $\vec{x}_b = \vec{x}^{opt}$, if $\rho > 0$
- **End while**

End TEA

Fig. 2. The Trusted Evolutionary Algorithm

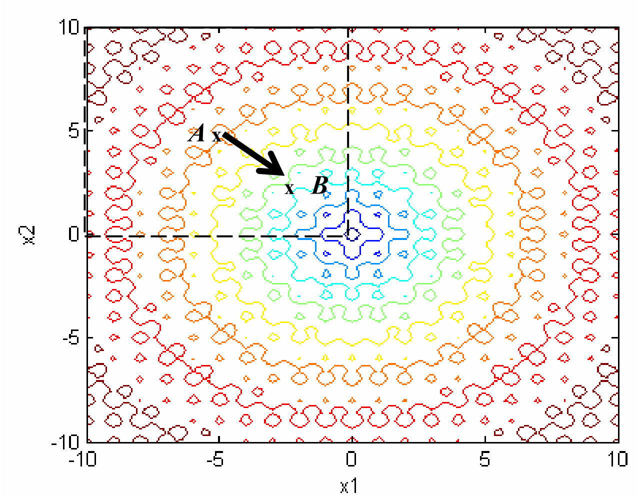
TEA begins its search using the canonical EA, i.e. EA with only exact function evaluations. During the canonical EA search, the exact fitness values obtained are archived in a central database together with the design vectors. After some initial search generations specified by the user, the algorithm proceeds by initializing the trust radius $\vec{\Delta} = \Delta_{1,}\Delta_{2,} ..., \Delta_n$. Here, the trust radius for dimension j is initialized as $\Delta_j = \min\left(x_{dj}^{\max} - x_{bj}, x_{bj} - x_{dj}^{\min}\right)$, where $x_{dj}^{\max}$ and $x_{dj}^{\min}$ are the maximum/minimum bounds of the m nearest points in the database to the current best point, $\vec{x}_b$. This makes good sense since the region in which the m points lie is regarded as the most trustworthy region where the search begins.

To provide a clear illustration on the search mechanisms of TEA, we use the Ackley benchmark function which is defined by:

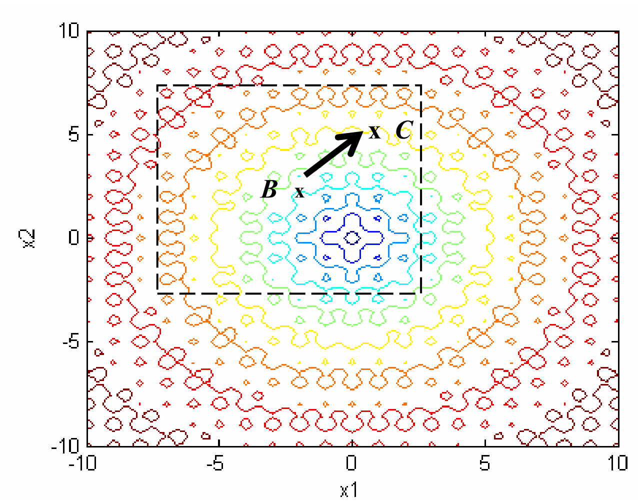
$$f(\vec{x}) = 20 + e - 20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos 2\pi x_i\right) \quad (7)$$

Figures 3(a)-(d) demonstrate the search process of TEA on the 2D Ackley function. The $\overline{EA}$ [1] search begins from the best solution provided by the canonical ES, $\vec{x}_b^1$, which is labeled in Figure 3(a) as point *A*. *A*fter *t* iterations, the $\overline{EA}$ using the surrogate model for fitness evaluations converges to new solution $\vec{x}_{opt}^1$ or point *B* in Figure 3(a). We consider the event of $\hat{f}(\vec{x}_{opt}^1) \leq \hat{f}(\vec{x}_b^1)$ for a minimization sub-problem and $f(\vec{x}_{opt}^1) \leq f(\vec{x}_b^1)$, i.e., surrogate is guiding the $\overline{EA}$ search in the correct direction leading to good search improvement, see Figure 3(a). Consider also that ρ>0.75, the trust radius $\vec{\Delta}^2$ for the next iteration remains unchanged since $\vec{x}_{opt}^1$ or point *B* falls within the boundary of the trusted region, i.e. $\| \vec{x}_{opt}^1 - \vec{x}_b^1 \|_\infty < \vec{\Delta}^1$. Starting from the newly found point B or $\vec{x}_b^2$, the $\overline{EA}$ search converges to point *C* or $\vec{x}_{opt}^2$, see Figure 3(b). However, since $f(\vec{x}_{opt}^2) > f(\vec{x}_b^2)$ and ρ<0, trust radius $\vec{\Delta}$ is thus reduced as the present surrogate displays low trustworthiness for the current search bound. Consequently, the search restarts from point *B* within the reduced $\vec{\Delta}$ as depicted in Figure 3(c). In the condensed region, $\overline{EA}$ converges subsequently to the new point *D* in Figure 3(d). Since the boundary of the trusted region is reached, i.e. $\| \vec{x}_{opt}^2 - \vec{x}_b^2 \|_\infty = \vec{\Delta}^2$ and considering that ρ>0.75, which suggests high trustworthiness in the surrogate to produce search improvements, $\vec{\Delta}$ is thus expanded in the next search iteration, i.e., $\vec{\Delta}^3 > \vec{\Delta}^2$. Thereafter, the search continues from point *D* and the process repeats until the termination conditions of the TEA are reached.

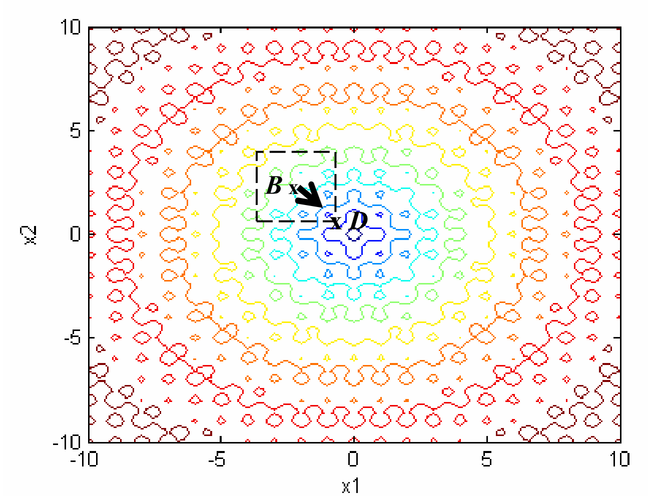
[1] $\overline{EA}$ refers to an EA that employs surrogate models for providing the fitness values on the population.



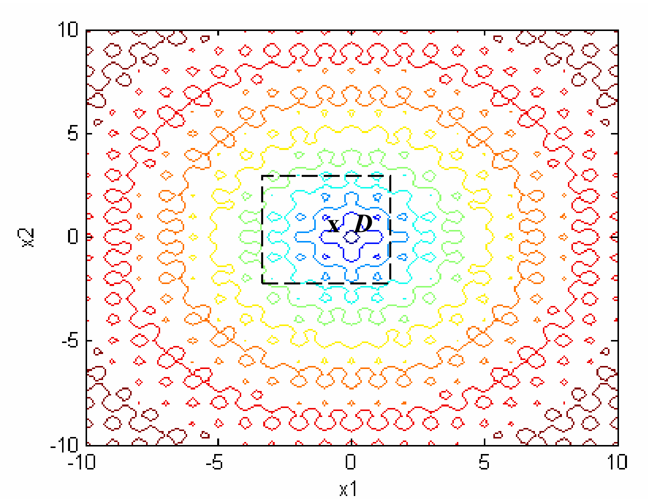
(a) Starting from initial guess $\vec{x}_b^1$ or point *A*, the $\overline{EA}$ search converges to $\vec{x}_{opt}^1$ or point B where $f(\vec{x}_{opt}^1) < f(\vec{x}_b^1)$. By assuming that ρ>0.75, $\vec{\Delta}$ is kept unchanged in the next iteration in the event that $\| \vec{x}_{opt}^1 - \vec{x}_b^1 \|_\infty < \vec{\Delta}^1$.



(b) Starting from initial guess $\vec{x}_b^2$ or B, the $\overline{EA}$ search converges to $\vec{x}_{opt}^2$ or point *C*. Since $f(\vec{x}_{opt}^2) > f(\vec{x}_b^2)$, ρ <0 and hence $\vec{\Delta}$ is reduced in the next iteration.



(c) Starting from initial guess $\vec{x}_b^2$ or point B, the $\overline{EA}$ search converges to $\vec{x}_{opt}^2$ or point *D* and $f(\vec{x}_{opt}^2) > f(\vec{x}_b^2)$. By assuming ρ>0.75, $\vec{\Delta}$ is increased in the next iteration.



(d) The $\overline{EA}$ search starts from initial guess $\vec{x}_b^3$ or point D with the increased trust radius.

Fig. 3. Search pattern of TEA on the 2D Ackley function.

IV. Empirical Study

In this section, we perform a numerical investigation of TEA and compare it to the canonical ES and original trust region line search framework. In particular, two multi-modal benchmark problems commonly used in the global optimization literature are adopted. The Ackley function is defined previously in equation (7) while the Griewank function is defined here as:

$$f(\vec{x}) = 1 + \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) \quad (8)$$

Here, the variable search bounds are set to [-10,10] and [-600,600] for the Ackley and Griewank functions, respectively. The global minimum for both functions is $f(\vec{x}) = 0.0$ and located at $x_i = 0.0$ for $i \in \{1,2,...,n\}$ where n is the dimensionality of the function.

In our experimental study, the initialization phase is conducted until a collection of 2*m* design points in the central database is reached. *m* is the number of training points used in the construction of the surrogates. Further, we consider radial basis neural network surrogates, which can approximate multiple-input multiple-output data efficiently, particularly when a few hundred data points are used for training. Let $\{\vec{x}_i, f(\vec{x}_i), i = 1,2,...,m\}$ denotes the training dataset, consists of *m* data pairs, where $\vec{x}_i$ is the input vector and $f(\vec{x}_i)$ is the corresponding output. Since, we are concerned with deterministic computer models, interpolating RBF approximation model of the following form is used.

$$\hat{f}(\vec{x}) = \sum_{i=1}^{m} \alpha_i K(\|\vec{x} - \vec{x}_i\|) \quad (9)$$

where $K(\|\vec{x} - \vec{x}_i\|) : \Re^d \to \Re$ is a radial basis kernel and $\alpha = \{\alpha_1, \alpha_2, ..., \alpha_m\}$ denotes the weight vector. Typical choices for the kernel include linear splines, thin-plate splines, cubic splines, Gaussian and multiquadric functions [24]. Here we consider the use of cubic splines for constructing surrogate models since earlier studies [27] suggests that both cubic and thin plate splines kernels are capable of providing models with good generalization capability at a low computational cost on high dimensional problems.

A. Comparison to Canonical ES

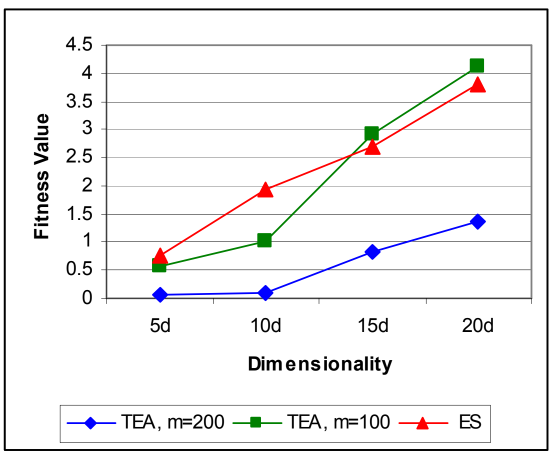
The average performance of the (1,10)-TEA and canonical (1,10)-ES from 10 independent runs for solving the benchmark test functions are summarized in Table 1 for dimensionalities *n*=5, 10, 15, and 20. These results are obtained based on the configurations of *m*=200, *t*=40, and for a maximum computational budget of 1000 exact function evaluations. Note that *t* is denoted here as the number of $\overline{EA}$ search iterations before the figure of merit, ρ, is measured to evaluate the trustworthiness of the trust radius and approximation model in generating fitness improvement (see Figure 2). The result reported in table 1 clearly indicates that TEA outperforms the canonical ES on both multi-modal benchmark problems for various dimensions under limited computational budget.

TABLE I
AVERAGE BEST FITNESS VALUES OBTAINED AT THE END OF 1000 EXACT FUNCTION EVALUATIONS FOR ACKLEY AND GRIEWANK FUNCTIONS FOR TEA AND ES

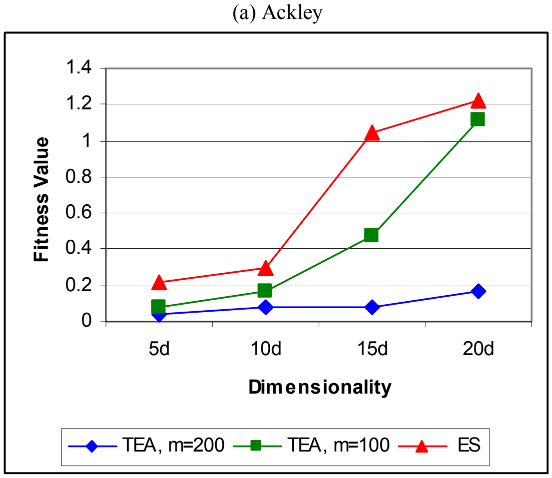
Function	Dimensionality (*n*)	TEA	ES
Ackley	5	0.0741	0.7453
	10	0.0945	1.9391
	15	0.8342	2.7053
	20	1.3703	3.8035

Griewank	5	0.0360	0.2165
	10	0.0751	0.2993
	15	0.0805	1.0443
	20	0.1686	1.2255

Further, we also study the effect of the additional parameters, i.e. m and t, on the performance of TEA. With t fixed at 40, we consider first the effect of training sample size. The averaged fitness values of (1,10)-ES and (1,10)-TEAs for m=100 and m=200 are summarized in Figure 4(a) and 4(b). From these figures, TEA is observed to fare poorer than the canonical ES on the 15D and 20D Ackley function when insufficient training points are employed, i.e. for m=100, resulting in inaccurate approximation models and poor search improvements. Note that the effect is more obvious on the higher dimensional problems, due to the effect of 'curse of dimensionality'. In contrast, TEA always outperforms the canonical ES in all parameter configurations considered on the Griewank function. In this case, the less accurate approximation for m=100 appears to generalize or smooth the multi-modality surface of the Griewank function pretty successfully, leading to excellent search performance in TEA.



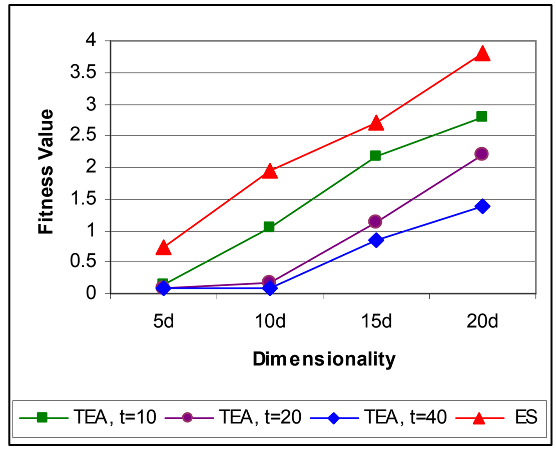
(a) Ackley



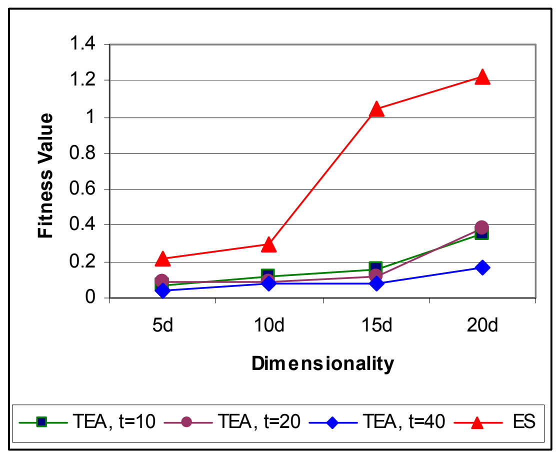
(b) Griewank

Fig. 4. Best fitness values obtained at the end of 1000 exact fitness evaluations for TEA m=100, 200, and canonical ES on: (a) Ackley, (b) Griewank benchmark functions.

In the same manner, we fixed m at 200 to identify some suitable values for t. The averaged fitness values of (1,10)-ES and (1,10)-TEAs for t=10, 20, 40, and m=200 are reported in Figure 5(a) and 5(b) when searching on the benchmark functions. From these results, TEA is shown to provide the best quality solution when t=40 and m=200 on both benchmark problems of diverse dimensionalities.



(a) Ackley



(b) Griewank

Fig. 5. Best fitness values obtained at the end of 1000 exact fitness evaluations for TEA t=10, 20, 40, and canonical ES on: (a) Ackley, (b) Griewank benchmark functions.

B. *Comparison to Original Trust Region Line Search Framework*

In this subsection, we evaluate the performance of the TEA against the original trust region framework for line search (TRF). For this purpose, we configure TEA for m=200, t=40, and computational budget of 1000 fitness evaluations. For fair comparison, TRF is configured to begin its search with a randomly generated database of 400 evaluated points, m=200 and maximum of (1000-400)=600 search iterations. Hence, the total computational budget is for both algorithms are 1000 fitness evaluations.

In the TRF, the gradient-based local solver is based on the Feasible Sequential Quadratic Programming (FSQP) [28] that employs cubic RBF models. The average search performance of (1,10)-TEA and TRF across 10 independent runs for solving two 20-dimensional benchmark problems are summarized in Table 2. It is shown from the results tabulated that TEA outperforms TRF significantly on the 20D Ackley function while slightly poorer on the 20D Griewank function, even when no form of gradient information have been used in the former. Further, the results in Table 2 also suggest that TRF is vulnerable to getting stuck at some local minima on the Ackley function, i.e., 5/10 independent runs in TRF, compared to only 1/10 in TEA, (see the italic values in Table 2).

TABLE 2
BEST FITNESS VALUES OBTAINED AT THE END OF 1000 EXACT FUNCTION EVALUATIONS FOR ACKLEY AND GRIEWANK FUNCTIONS FOR TEA AND TRF

Trial	Ackley 20D TEA	Ackley 20D TRF	Griewank 20D TEA	Griewank 20D TRF
1	0.4172	2.9125	0.0321	0.0005
2	1.4827	*3.1694*	0.0401	0.0014
3	0.8310	*8.5273*	0.0235	0.2761
4	2.1603	*4.2022*	0.1281	0.0012
5	0.7056	2.6362	0.2766	8.37×10^{-5}
6	0.7352	*9.0514*	0.1692	6.61×10^{-5}
7	0.6281	2.6354	0.1562	0.0006
8	*4.6948*	*3.9832*	0.0255	0.2845
9	1.3253	*3.4658*	0.0025	0.5358
10	0.7224	2.4438	0.8316	0.1170
Mean (μ)	1.3703	4.3027	0.1686	0.1217
Standard deviation (σ)	1.2127	2.3117	0.2360	0.1760

V. CONCLUSION

In this paper, a trust-region inspired evolutionary search or the Trusted Evolutionary Algorithm (TEA) is presented. In contrast to earlier work, the focus of TEA is on predicting search improvement in the context of optimization as opposed to the quality of the approximation, which is a treated here as a secondary objective. Empirical study conducted using two benchmark functions shows that the proposed TEA converges to better solution than conventional evolutionary search under a limited computational budget. In addition, it is worth noting that the present work has considered only the use of zero order approximation models. In our future work, novel TEA that employs more sophisticated models, for example, first and second order models, will be considered.

VI. ACKNOWLEDGEMENT

D. Lim and Y.S. Ong would like to thank Honda Research Institute Germany for funding this research work, and members of Nanyang Technological University for providing computing support and resources to make this work possible.

REFERENCES

[1] M.K. Karakasis, A.P. Giotis, and K.C. Giannakoglou, "Inexact information aided, low-cost, distributed genetic algorithms for aerodynamic shape optimization", *International Journal for Numerical Methods in Fluids*, vol. 43, pp. 1149–1166, 2003.

[2] P. Hajela, J. Lee, " Genetic algorithms in multidisciplinary rotor blade design", In *Proceedings of 36th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Material Conference*, New Orleans, pp. 2187-2197, 1995.

[3] H. Fang, P. Ross, and D. Corne, "A Promising Genetic Algorithm Approach to Job-shop Scheduling, Rescheduling and open-shop scheduling problems", *Proceedings of the 5th International Conference on Genetic Algorithms*, ICGA-93. Morgan Kaufmann, 1993.

[4] C.G. Johnson and J.J.R. Caldalda, "Introduction: Genetic Algorithms in Visual Art and Music", *Leonardo* - Volume 35, Number 2, April 2002, pp. 175-184.

[5] P.B. Nair, A.J. Keane, "Passive Vibration Suppression of Flexible Space Structures via Optimal Geometric Redesign", *AIAA Journal* 39(7), pp. 1338-1346, 2001.

[6] F.S. Lesh, "Multi-dimensional least-square polynomial curve fitting", *Communications of ACM*, vol. 2, no. 9, pp. 29-30, 1959.

[7] Z. Zhou, Y.S. Ong, M.H. Nguyen, D. Lim, "A Study on Polynomial Regression and Gaussian Process Global Surrogate Model in Hierarchical Surrogate-Assisted Evolutionary Algorithm", *IEEE Congress on Evolutionary Computation*, Edinburgh, United Kingdom, September 2-5, 2005.

[8] A.A. Guinta and L. Watson, "A comparison of approximation modelling techniques: polynomial versus interpolating models", in *Proc. the 7th AIAA/ USAF/ NASA/ ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA-98-4758*, St. Louis, MO, USA, Sept. 1998.

[9] L. Gräning, Y. Jin, B. Sendhoff, "Efficient Evolutionary Optimization Using Individual-based Evolution Control and Neural Networks: A Comparative Study", *Proceeding of European Symposium on Artificial Neural Networks (ESANN)*, Bruges (Belgium), 27-29 April 2005.

[10] L. Willmes, T. Bäck, Y. Jin, and B. Sendhoff, "Comparing Neural Networks and Kriging for Fitness Approximation in Evolutionary Optimization", In *Proceedings of IEEE Congress on Evolutionary Computation*, pages 663--670, 2003.

[11] A. Ratle, "Kriging as a surrogate fitness landscape in evolutionary optimization", *Artificial Intelligence for Engineering Design Analysis and Manufacturing*, vol. 15, no. 1, pp. 37-49, 2001.

[12] D. Büche, N.N. Schraudolph, and P. Koumoutsakos, "Accelerating evolutionary algorithms with gaussian process fitness function models", *IEEE Transactions on Systems, Man, and Cybernetics* part C, 2004.

[13] Z. Zhou, Y.S. Ong, P.B. Nair, A.J. Keane, and K.Y. Lum, "Combining Global and Local Surrogate Models to Accelerate Evolutionary Optimization", *IEEE Transactions on Systems, Man and Cybernetics (SMC), part C*, In press, expected July 2006.

[14] M.A. El-Beltagy, P.B. Nair, A.J. Keane, "Metamodelling Techniques For Evolutionary Optimization of Computationally Expensive Problems: Promise and Limitations", *Proceeding of the Genetic and Evolutionary Computation Conference*, 1999.

[15] Y. Jin, M. Olhofer, B. Sendhoff, "A Framework for Evolutionary Optimization with Approximate Fitness Function", *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 5, October 2002.

[16] H. Ulmer, F. Streichert and A. Zell, "Model-assisted Evolution Strategies", In Y. Jin, editor, *Knowledge Incorporation in Evolutionary Computation*, pages 333--357. Springer, Berlin, 2004.

[17] H. Ulmer, F. Streichert, and A. Zell, “Evolution strategies assisted by Gaussian processes with improved pre-selection criterion”, *Proceedings of IEEE Congress on Evolutionary Computation*, pp. 692-699, 2003.

[18] Y. Jin and B. Sendhoff, “Fitness approximation in evolutionary computation: A survey”, *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1105-1112, 2002.

[19] J.E. Dennis, V. Torczon, “Managing Approximation Models in Optimization and Nonlinear Equations”, Prentice Hall, Anglewood Cliffs, NJ, 1983.

[20] N.M. Alexandrov, J.E. Dennis, R.M. Lewis, and V. Torczon, “A Trust Region Framework for Managing Use of Approximation Models in Optimization”, *Journal on Structural Optimization*, Vol. 15, No. 1, pp. 16-23, 1998.

[21] Y.S. Ong, P.B. Nair, and A.J. Keane, "Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modeling", *AIAA Journal*, Vol. 41, No. 4, pp 687-696, 2003.

[22] Y.S. Ong, P.B. Nair, K.Y. Lum, “Evolutionary Algorithm with Hermite Radial Basis Function Interpolants for Computationally Expensive Adjoint Solvers”, *Computational Optimization and Applications*, conditionally accepted, 2005.

[23] J. Morè, “Recent Developments in Algorithms and Software for Trust Region Methods”, *Mathematical Programming. The State of Art*, Bonn 1982, Bachem A., Grotschel, and Korte B., eds., Springer-Verlag, Berlin, 1983, pp. 258-287.

[24] V. Vapnik, “Statistical Learning Theory”, John Wiley and Sons, NY, 1998.

[25] C. Bishop, “Neural Networks for Pattern Recognition”, Oxford University Press, 1995.

[26] K.C. Giannakoglou, “Design of Optimal Aerodynamic Shapes Using Stochastic Optimization Methods and Computational Intelligence”, *International Review Journal Progress in Aerospace Sciences,* vol. 38, No. 5, pp. 43-76, 2002.

[27] H.M. Gutmann, “On the Semi-norm of Radial Basis Function interpolants”, Dept. Applied Math. Theor. Phy., Univ. Cambridge, UK., Tech. Rep. DAMTP 2000/NA04, 2000.

[28] C.T. Lawrence and A.L. Tits, “A Computationally Efficient Feasible Sequential Quadratic Programming Algorithm”, *Society for Industrial and Applied Mathematics*, Vol. 11, No. 4, pp. 1092-1118, 2001.