

Evaluation of User Fatigue Reduction Through IEC Rating-Scale Mapping

Wang, Shangfei

Department of Computer Science, University of Science and Technology of China

Takagi, Hideyuki

Faculty of Design, Kyushu University

<https://hdl.handle.net/2324/1808896>

出版情報 : IEEE International Workshop on Soft Computing as Transdisciplinary Science and Technology (WSTST). 4, pp.672-681, 2005-05-25. IEEE

バージョン :

権利関係 :

Evaluation of User Fatigue Reduction Through IEC Rating-Scale Mapping

Shangfei Wang¹

Hideyuki Takagi²

sfwang@ustc.edu.cn, takagi@design.kyushu-u.ac.jp

Faculty of Design, Kyushu University, Fukuoka, 815-8540, Japan

¹ (On leave from the Department of Computer Science, University of Science and Technology of China)

Abstract:

We evaluate the convergence speed of an Interactive Evolutionary Computation (IEC) using a rating-scale mapping for user fatigue reduction. First, we introduce the concept of mapping users' relative ratings to an "absolute scale"; this allows us to improve the performance of the IEC subjective evaluation characteristic predictor, which can in turn accelerate EC convergence and reduce user fatigue. Second, we experimentally evaluate the effectiveness of the proposed method using seven benchmark functions instead of a human user. The experimental results show that the convergence speed of an IEC using the proposed absolute rating data-trained predictor is much faster than an IEC using a conventional predictor trained using relative rating data.

1 Introduction

Interactive Evolutionary Computation (IEC) is an optimization technology that adopts evolutionary computation (EC) for system optimization based on subjective human evaluation [1]. Simply put, it is an EC technique where the fitness function is replaced with a human user. In this sense, IEC is a technology that embeds human preference, intuition, emotion, and psychological aspects into the target system. Over the past 10 years, the number of papers on IEC has increased and its areas of application have expanded to cover a wide variety of fields, such as computer graphics, music, artistic design, signal processing, data mining, virtual reality, micro-electrical mechanical system design, geophysical simulation and others [1].

However, because IEC users must evaluate every individual of every generation, user fatigue is its biggest technological problem. Fatigue is especially serious when the population size and number of generations is large. To make IEC technology more practical, we must improve the algorithm to compensate for this.

There are three approaches for reducing user fatigue: improving the input interface, improving the display interface, and accelerating EC convergence [1]. Predicting IEC users' evaluation characteristics is one way of accelerating EC

convergence [2]. If an IEC has a predictive function as shown in Fig. 1, a normal EC search with a large population size can be used behind IEC user' evaluation [3], allowing the IEC display interface [2] and searching capability to be improved, thus accelerating convergence and reducing user fatigue.

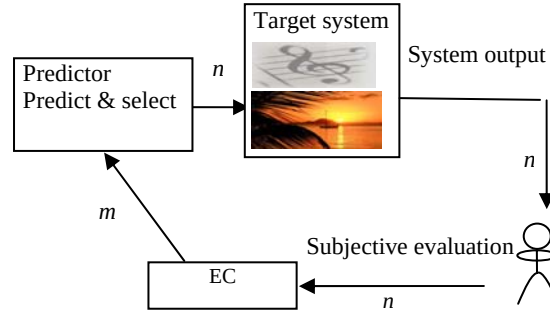


Fig. 1. IEC system with a conventional predictor ($m \gg n$).

There are remaining unexplored methods for predicting user evaluation characteristics. An IEC user's rating-scale is not absolute over all generations but rather relative within each generation, i.e. the rating values from generations long past and from recent generations may be different even if the evaluation targets are the same. These generation-to-generation differences appear as noise to the algorithm which learns and predicts the users' evaluation characteristics.

To overcome this shortcoming, we proposed the concept of *absolute scale* and define a method for mapping users' relative evaluation data to the absolute scale [3]. With this concept, only the absolute rating data is used to learn users' characteristics. Since the noisiness of data is reduced, the performance of the user evaluation characteristic predictor should be improved, and an acceleration of EC convergence is expected. This should, in turn, result in less fatigue for the users during their interaction with the system. The effectiveness of our approach is evaluated through simulation tests using seven benchmark functions instead of a real human IEC user whose evaluation fluctuates and takes time.

2 IEC with absolute rating-scale mapping

Using relative rating values is easier and less stressful for IEC users, but it causes poor learning because the rating values from long past generations and neighboring generations may be different even if the evaluation target is the same. For example, the best individual in the i -th generation may be worse than the worst individual in the $(i+j)^{\text{th}}$ generation, as shown in Fig. 2. To the algorithm which learns the users' evaluation characteristics, these differences appear as noise. To reduce this noise, we propose a mapping from the relative rating values to an absolute scale [2].

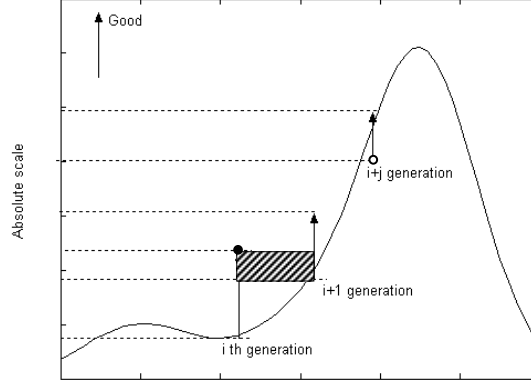


Fig. 2. Relationship between a relative scale and an absolute scale. The open circle ($\circ$) indicates the worst individual in the $(i+j)^{\text{th}}$ generation, and closed circle ($\bullet$) indicates the best individual in the i -th generation. The area enclosed by the bias line indicates the l individuals in the i -th generation that remained in the next generation.

Given a set of individuals in the i -th generation $\{x_j^i, j = 1, \dots, n\}$, a user's evaluation values for these individuals are $\{f_j^i, j = 1, \dots, n\}$. Because a user's evaluation is relative for each generation, the l individuals $\{x_j^i, j = 1, \dots, l\}$ from the i -th generation that remain in the next generation will have different evaluation values $\{f_j^{i+1}, j = 1, \dots, l\}$ in the $(i+1)$ generation. The relative evaluation values are thus mapped to absolute values, F , according to Eq. (1).

$$\begin{aligned} F_j^{i+1} &= f_j^{i+1} - \frac{1}{l} \sum_{j=1}^l (f_j^{i+1} - f_j^i) \quad j = 1, \dots, n \\ F_j^i &= f_j^i \end{aligned} \quad (1)$$

This means that we use the average difference in evaluation between two neighboring generations to shift the relative scale to an absolute one. After mapping, the absolute evaluation data are used to learn the user's evaluation characteristic. Once a system learns the characteristic, it can be used as a fitness function in an EC that works behind the scenes to search for the global optimum using the user evaluations and ratings. We call this learning system a predictor for human evaluation that works as a pseudo-IEC user in this paper.

Fig. 3 shows the structure of an IEC with our proposed predictor using an absolute scale. In this paper, we use a neural network (NN) as the predictor, but the predictor need not be restricted to NN. In each generation, the NN predictor is trained using EC individuals and the mapped absolute values of the IEC user's subjective fitness values for the individuals in all past generations. The trained NN

predictor takes as inputs m individuals generated by EC and outputs the predicted absolute fitness values of the IEC user for those individuals. Of these m individuals, only the best n individuals are actually displayed to the IEC user, which is a benchmark function in Section 3 in the next generation.

Although the mapping method described in this section is not specific, it is easy to implement and it can reduce the noise caused by relative evaluation. We expect that an IEC using the proposed predictor will converge more quickly than one which uses a conventional predictor, and evaluate such a system's performance in Section 3.

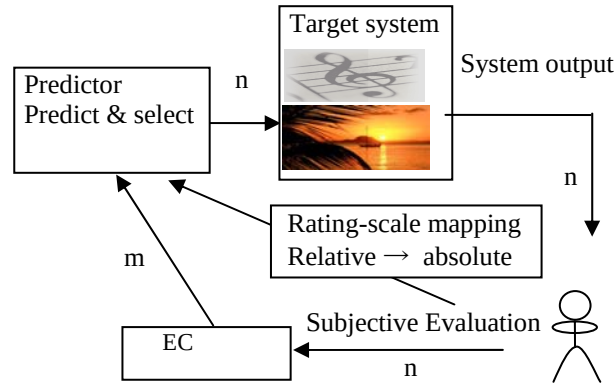


Fig. 3. IEC system with the proposed predictor ($m \gg n$).

3 Experimental evaluations

3.1 Experiment conditions

The purpose of our experiment is to evaluate how much the use of the proposed absolute scale accelerates the convergence of the EC search. The final validation of the system should be conducted using subjective tests of an IEC with the proposed predictor. However, since IEC deals with subjective evaluation values that depend on the application task and the subject's perceived value of the task, in our preliminary investigation we evaluate the effect of the proposed method with a simulated subjective evaluation for the task given in this section.

The task is to find x_i at the global optimum of seven benchmark functions (see Table 1). A genetic algorithm (GA) is used as one of the EC technologies. To simulate a human's relative evaluation, we first calculate the function values of each of the individuals in a generation according to the equations in the Table 1, and then change these values to a scale with five degrees using uniform quantization. These five degrees correspond to a human's relative evaluation.

The performance measure is minimum value versus computation time. We compared three methods: an interactive genetic algorithm (IGA) with the proposed predictor using an absolute scale, an IGA with a conventional predictor using a

relative scale, and an IGA without a predictor. Table 2 shows the parameter settings of the IGA in the experiments.

The NN predictor inputs 200 ($m = 200$) individuals, selects the best 20 ($n = 20$) individuals, and passes them on to the target system. See m and n in Fig.3.

Table 1. Experimental benchmark function.

DeJong's Function 1-D	$f(x) = \sum_{i=1}^n x_i^2 \quad -512 \leq x_i \leq 512$
Rosenbrock's Function	$f(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 - (1 - x_i)^2 \quad -2.048 \leq x_i \leq 2.048$
Rastrigin's Function	$f(x) = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)) \quad -5.12 \leq x_i \leq 5.12$
Schwefel's Function	$f(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i }) \quad -500 \leq x_i \leq 500$
Griewangk's Function	$f(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad -600 \leq x_i \leq 600$
Hartman's Function	$f(x) = -\sum_{i=1}^4 c_i \exp\left[-\sum_{j=1}^n a_{ij} (x_j - p_{ij})^2\right] \quad 0 \leq x_j \leq 1$
Sum of Different Powers Function	$f(x) = \sum_{i=1}^n x_i ^{(i+1)} \quad -1 \leq x_i \leq 1$

Table 2. Experimental conditions of the GA.

Parameters	Value
Population size	20
Crossover rate	0.7
Mutation rate	0.01
No. of generations	20
GA coding	Binary coding
Bit length	20

3.2 Experiment results

Firstly, we evaluated the effectiveness of mapping the relative ratings to an absolute scale. Fig. 4 is a sample result when a 1-D DeJong's function is used. The figure on the left shows absolute fitness versus x , while the right figure shows relative fitness versus x . From the Fig. 4, it can be seen that the shape using absolute fitness is much more similar to shape of a 1-D DeJong's function than that which is obtained using relative fitness. This means that this kind of scale-mapping can reduce the noise in the learning data provided to the predictor.

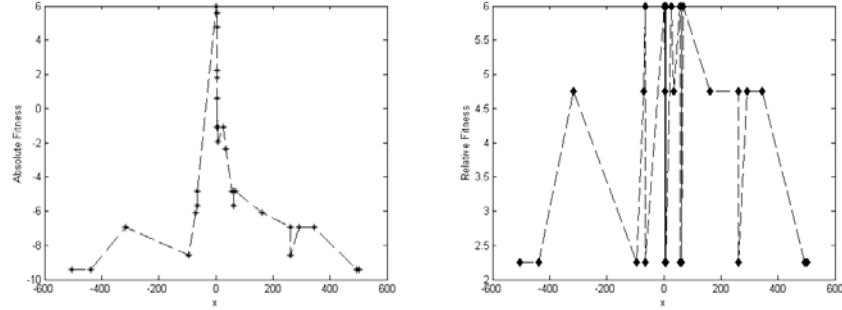


Fig. 4. The effective of rating-scale mapping from relative (right) to absolute one (left) for a 1-dimensional DeJong' function.

The effectiveness of an IEC using our proposed predictor was estimated using seven benchmark functions. Fig. 5 shows the experimental results. It is clear that the IGA with the proposed predictor using an absolute scale converges faster than not only the normal IGA without a predictor, but also the IGA using a conventional predictor with a relative scale.

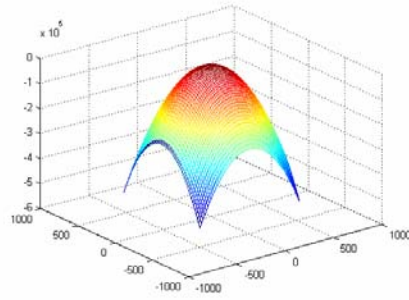


Fig. 5-1.a Geometric characteristic of DeJong's function 1.

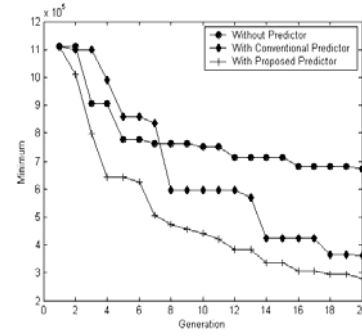


Fig. 5-1.b Convergence characteristics comparison for DeJong's function 1.

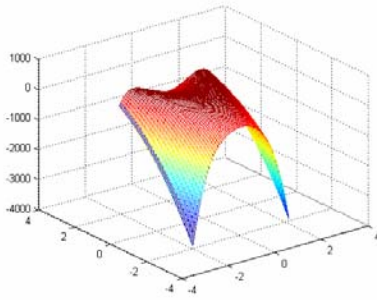


Fig. 5-2.a Geometric characteristic of Rosenbrock's function.

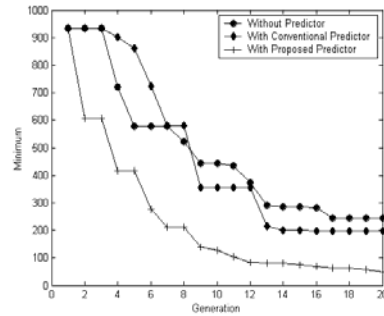


Fig. 5-2.b Convergence characteristics comparison for Rosenbrock's function.

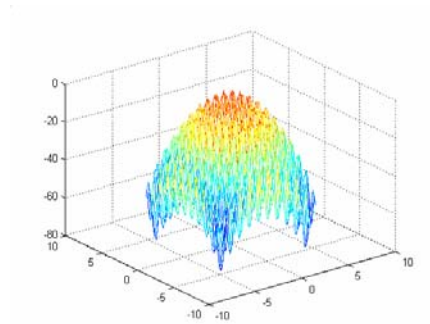


Fig. 5-3.a Geometric characteristic of Rastrigin's function.

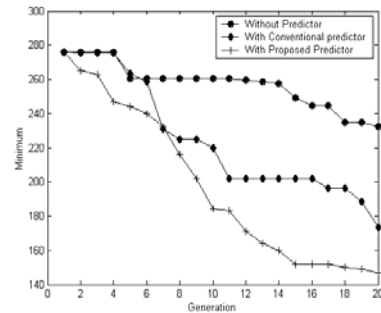


Fig. 5-3.b Convergence characteristics comparison for Rastrigin's function.

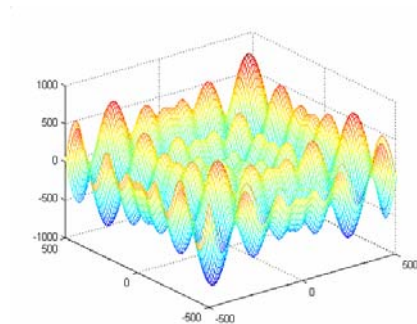


Fig.5-4.a Geometric characteristic of Schwefel's function.

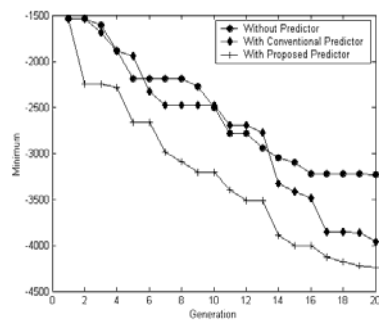


Fig. 5-4.b Convergence characteristics comparison for Schwefel's function.

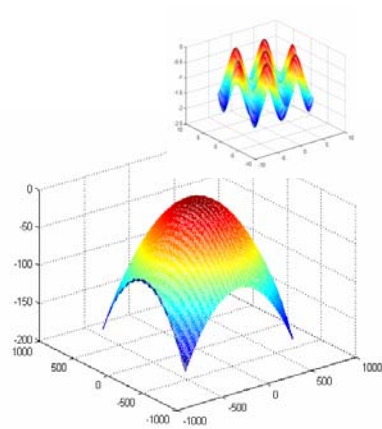


Fig. 5-5.a Geometric characteristic of Griewangk's function. The shape of the global area differs from DeJong's function 1

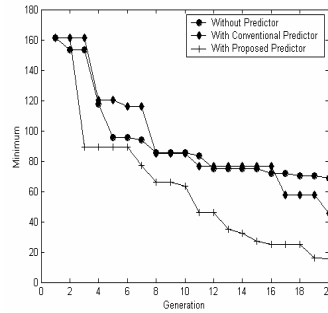


Fig. 5-5.b Convergence characteristics comparison for Griewangk's function.

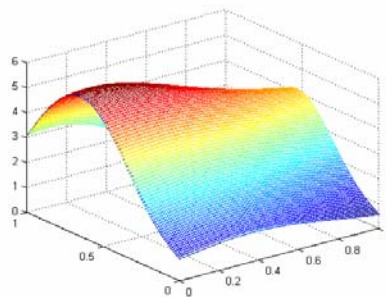


Fig. 5-6.a Geometric characteristic of Hartman's function.

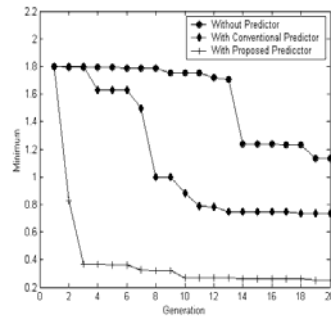


Fig. 5-6.b Convergence characteristics comparison for Hartman's function.

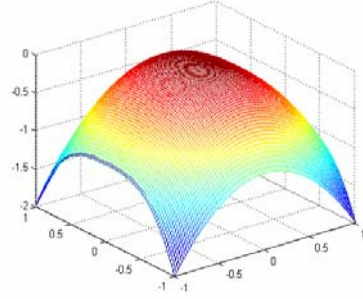


Fig. 5-7.a Geometric characteristic of Sun of Different Powers Function.

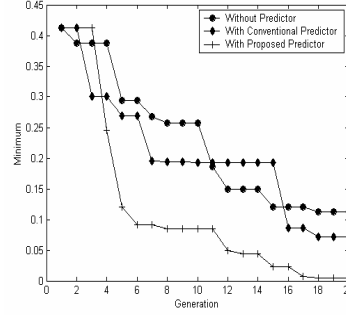


Fig. 5-7.b Convergence characteristics comparison for Sun of Different Powers Function.

Fig.5. Convergence characteristics comparison for normal IGA without a predictor, IGA with a conventional predictor using a relative scale, and IGA with the proposed predictor using an absolute scale.

4 Discussions

The experimental results obtained using the seven benchmark functions in Section 3 showed that our proposed method, using an NN predictor trained by estimated absolute rating values, accelerated the convergence of the global optimum search in all cases. In real IEC tasks, this acceleration in early generations will result in reduced IEC user fatigue.

It is expected that this acceleration of convergence is realizable not only for the seven benchmark functions, but also in general cases. The main reason why a significant acceleration of convergence was obtained in Section 3 is that the noise in the NN predictor training data is reduced. This is easily understood from the description of the method in the first paragraph in Section 2 and the simple experiment illustrated in Fig. 4. The noise reduction is not limited to the function in Fig. 4, but will be realized for all the benchmarks.

The experimental results in Fig. 5 show that a normal EC search with many individuals and a predictor, regardless of whether it uses relative or absolute fitness values, is effective for accelerating IEC convergence.

We therefore expect that our proposed introduction of an absolute scale-based predictor for the absolute values of the IEC user's subjective fitness values will accelerate IEC convergence in general.

5 Conclusions

We introduced the idea of using an *absolute scale* for reducing the noise in training data used with predictors that learn IEC users' subjective evaluation characteristics. We also proposed a concrete method of mapping relative data to the absolute scale. Simulation results have showed that scale-mapping can reduce the noise in learning data used by the predictor. The proposed predictor using absolute evaluation scores therefore has better prediction performance than conventional predictors, resulting in faster convergence in early generations. We discussed why we believe this effect can be expected to work in general. This convergence in early generations will reduce the problem of IEC user fatigue, which is our objective.

The next step in our evaluation of this technique will be to apply the proposed approach to some real tasks and evaluate its effectiveness in reducing IEC user fatigue with subjective tests.

Acknowledgments

This paper was supported by 21st Century COE program, "Design of Artificial Environments on the Basis of Human Sensibility."

Reference

- [1] Takagi, H. (2001), "Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation," Proceedings of the IEEE, vol. 89, no. 9, pp.1275-1296.
- [2] Ohsaki, M. and Takagi, H. (1998), "Improvement of Presenting Interface by Predicting the Evaluation Order to Reduce the Burden of Human Interactive EC Operations," IEEE Int. Conf. on System, Man, and Cybernetics (SMC1998), pp.1284-1289.
- [3] Wang, S. F. and Takagi, H. (2005), "Improving the Performance of Predicting Users' Subjective Evaluation Characteristics to Reduce Their Fatigue in IEC," J. of Physiological Anthropology Applied Human Science, vol. 24, no. 1, pp.121-125.