Correspondence

Neighborhood Topologies in Fully Informed and Best-of-Neighborhood Particle Swarms

James Kennedy and Rui Mendes

***Abstract*—In this study, we vary the way an individual in the particle swarm interacts with its neighbors. The performance of an indiviual depends on population topology as well as algorithm version. It appears that a fully informed particle swarm is more susceptible to alterations in the topology, but with a good topology, it can outperform the canonical version.**

***Index Terms*—Algorithms, optimization, particle swarm.**

I. Introduction

The particle swarm algorithm is based on a social-psychological model of social influence and social learning. A population of candidate problem solutions, randomly initialized in a high-dimensional search space, discovers optimal regions of the space through a process of individuals' emulation of the successes of their neighbors. This correspondence investigates an alternative implementation of social neighborhoods within the particle swarm framework.

In the traditional particle swarm, each individual has some number of neighbors, with a mutual influence between them. On each iteration of the program loop, the individual queries its neighbors to determine which one has had the best success with the problem thus far and uses the location of that success as well as the location of its own previous best success to choose a new point in the search space to test.

This represents an oversimplification of the social-psychological view that individuals are more affected by those who are more successful, persuasive, or otherwise prestigious. In the human society, it is more accurate to say that the social neighborhood provides a wealth of possible models whose behavior may be emulated, and individuals seem to be affected by some kind of statistical summary of the state of their immediate social network rather than the unique performance of one individual (e.g., [2] and [7]). This correspondence reports on the research with a version of particle swarm where an individual is influenced by the success of all of its neighbors, rather than just the best one. The fully informed particle swarm (FIPS) is compared with the canonical version, using a variety of neighborhood structures.

II. Neighborhoods

The earliest reported particle swarm version [3], [4] used a kind of topology that is known as *gbest*. The source of social influence on each particle was the best performing individual in the entire population. This is equivalent to a sociogram or social network where every individual is connected to every other one. The *gbest* topology was acceptable for the first applications, which typically involved finding a matrix of weights for a feedforward neural network. The function landscape of this kind of problem is largely made up of long gradients, where the problem is, first, to find the best gradient region of the search space and, second, to find the minimum of that region.

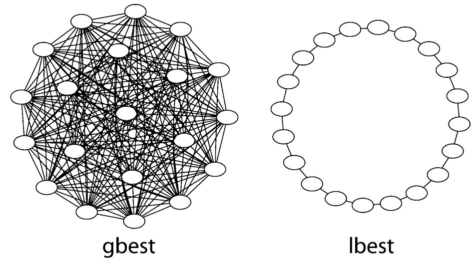

Fig. 1. *Gbest* and *lbest* population sociometries.

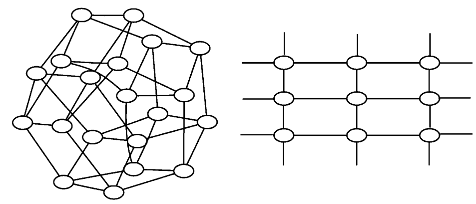

Fig. 2. von Neumann or square topology resulted in a better performance in the canonical particle swarm than either of the traditional topologies. It is formed by arranging the population in a grid and connecting neighbors above, below, and to the right and left. The edges of the matrix are wrapped. A population of 20 individuals is shown on the left, and a local region is on the right.

Many problems contain cliffs, variable interactions, and other features that are not typified by smooth gradients and need a more robust algorithm. The *lbest* topology (Fig. 1) was proposed as a way to deal with more difficult problems. In the *lbest* sociometric structure, each individual is connected to—that is, it influences and is influenced by—its immediate neighbors in the population array. The fifth individual in the population, for instance, interacts with the fourth and the sixth ones. The individuals are typically included in their own neighborhoods to be the sources of influence upon themselves, if they have found the best problem solution in the neighborhood so far.

The *lbest* topology offered the advantage that subpopulations could search diverse regions of the problem space. As some individuals in one part of the population influenced one another to focus on one local optimum, the other part of the population could search around another. The flow of information from one part of the population to another was moderated by the necessity of "persuading" intermediate individuals to search in a particular area. Once individuals gave their best performance in that region, they could influence their neighbors, and the influence could slowly travel through the neighborhood.

Kennedy and Mendes [6] tested some neighborhood configurations and discovered some, for instance, the "von Neumann" neighborhood, that performed better than the standard ones on a suite of standard test problems. The von Neumann neighborhood, named so after its use in cellular automata was pioneered by John von Neumann (Fig. 2), comprises a kind of square: The population is arranged in a rectangular matrix, for instance, 5×4 in a population of 20 individuals, and each individual is connected to the individuals above, below, and on both of its sides, wrapping the edges.

The real research question was, and continues to be, the discovery of aspects of the sociometry that actually affect the population's ability to optimize functions. Kennedy and Mendes found great variation among

Manuscript received May 9, 2003; revised November 29, 2005.

J. Kennedy is with the U.S. Bureau of Labor Statistics, Washington, DC 20212 USA (e-mail: Kennedy_Jim@bls.gov).

R. Mendes is with the Department of Informatics, University of Minho, 4704-553 Braga, Portugal (e-mail: azuki@di.uminho.pt).

Digital Object Identifier 10.1109/TSMCC.2006.875410

TABLE I
FREQUENCIES OF SOCIOMETRIES IN EACH DEGREE × CLUSTERING CONDITION

	Clustering								
k	0	1	2	3	4	6	7	8	Total
3	102	129	0	0	0	0	0	0	231
5	59	128	60	64	0	0	0	0	311
10	21	74	212	122	172	34	106	60	801

a small number of topologies, but it was not clear exactly how the topology or neighborhood structure interacted with the features of the function. Further, it seemed clear that some topologies were simply superior for solving the problems featured in the test suite, while some were good for nothing.

III. Canonical Particle Swarm and FIPS

This study was carried out with canonical best-neighbor particle swarms. The model can be summarized by a pair of formulas: the first modifying the particle's velocity and the second modifying its position in the parameter space

$$\boldsymbol{v}_i \leftarrow \chi\left(\boldsymbol{v}_i + U\left(0,\frac{\varphi}{2}\right)\cdot(\boldsymbol{p}_i - \boldsymbol{x}_i) + U\left(0,\frac{\varphi}{2}\right)\cdot(\boldsymbol{p}_g - \boldsymbol{x}_i)\right)$$
$$\boldsymbol{x}_i \leftarrow \boldsymbol{x}_i + \boldsymbol{v}_i$$

where $\boldsymbol{v}_i$ is the velocity vector for individual i; χ is a constriction coefficient that causes convergence of the individual trajectory in the search space and whose value is typically approximately 0.729 [1]; U (min, max) is a uniform random number generator; φ is the acceleration constant, which helps control the convergence of particles and is set to 4.1 following Clerc's analysis [1]; $\boldsymbol{p}_i$ is the best location found so far by particle i; $\boldsymbol{x}_i$ is the current position; and the index g takes the value of the index i's best neighbor. The result of iterating these two formulas is a kind of oscillatory trajectory around the average of i's and g's best points so far. In the course of the oscillations, the particle is likely to encounter a better position than its previous best, resulting in updating the $\boldsymbol{p}_i$. If i is the best in its neighbor's neighborhood, then that neighbor will be influenced to oscillate around a point defined in part by i's history, in its turn.

The FIPS model can be considered a generalization of the canonical one. While the canonical model adds two terms to the velocity and divides the constant φ in half to weight each of them, the FIPS distributes the weight of φ across the entire neighborhood

$$\boldsymbol{v}_i \leftarrow \chi\left(\boldsymbol{v}_i + \sum_{n=1}^{N_i}\frac{U(0,\varphi)(\boldsymbol{p}_{\mathrm{nbr}(n)} - \boldsymbol{x}_i)}{N_i}\right)$$
$$\boldsymbol{x}_i \leftarrow \boldsymbol{x}_i + \boldsymbol{v}_i$$

where N_i is the number of neighbors particle i has and $\mathrm{nbr}(n)$ is i's nth neighbor. If $\mathrm{nbr}(n)$ includes only i itself and its best neighbor, then this formula degenerates to the canonical version. In the results reported in this correspondence, various neighborhood topologies will be considered for both the canonical and FIPS algorithms.

IV. Neighborhoods in the Two Versions

In the canonical particle swarm, where only the best member of the neighborhood has any effect, it appears that the size of the neighborhood will directly affect the quality of influence. A bigger neighborhood has a better chance of containing a relatively excellent solution. In comparisons of the *lbest* and *gbest* canonical algorithms, a difference has been noted in their convergence speed and the ability to search over local optima [5]. The *gbest* topology (i.e., the biggest neighborhood possible) has often been shown to converge on optima more quickly than *lbest*, but *gbest* is also more susceptible to the attraction of local optima since the population rushes unanimously toward the first good solution found. One way of understanding this is to consider the success of simulated annealing, where poorer candidate solutions are sometimes retained: the *lbest* population allows particles to influence their neighbors even when they are poorer than the population best and this influence of weaker individuals results in a stronger outcome.

A larger neighborhood is likely to contain better solutions. In the canonical particle swarm, where only the best neighbor has any effect, this means that the influence on the particle tends to be of a higher quality, which is not always good, since it promotes premature convergence. In the FIPS, a nearly opposite phenomenon is expected: A larger neighborhood is more likely to contain individuals from diverse regions of the problem space and the target particle will be pulled toward an average that is in between optima. The result should be confusion in the presence of many influences.

V. Method

The population topologies were created from random graphs, optimized to meet as nearly as possible some criteria. All the graphs had 20 nodes and were connected, which means that there was at least one path from every node to every other one. Three levels of mean degree were investigated, and the nodes had, on average, three, five, or ten neighbors. Furthermore, each graph had high or low clustering, defined as the mean number of i's neighbors that were neighbors to one another. The variance of k (mean degree) and clustering were also manipulated and so the graphs had high or low variance in these features. The graphs were fit to the criteria by using a simulated annealing method.

As it was not equally easy to make graphs that met the various criteria, 1343 population structures were made in the end. The distribution of the test conditions is shown in Table I.

In this study, the most important independent variable was the algorithm type, that is, canonical particle swarm versus FIPS. Analyses of the minor independent variables have shed light mainly on the differences found between the two versions. All conditions were tested on a standard suite of test functions.

Two dependent variables were examined. The first was the mean best performance of a population after 1000 iterations of the algorithm. The results of the six functions were combined by standardizing results and then averaging. First, all the results for a function were transformed to have a mean of 0.0 and a standard deviation of 1.0. Then, records containing one result on each function were created by combining the first results for each function into one record, then the second results into a second record, and so on. The standardized results were then averaged for each record. Standardizing and combining results gives us a number that indicates how well the topology performed across all the functions: We are not interested in finding an algorithm that can solve one kind of problem and not others. Most of the results reported in this paper use averages of all records for a population topology, i.e., one value for each sociogram.

This is an important measure of how well the algorithm is able to find a good point on the search space, but does not indicate an ability to find the global optimum since it is possible that a good result is

TABLE II
PARAMETERS AND CRITERIA FOR THE SIX TEST FUNCTION CONDITIONS

Function	Dimensions	Initial range	Criterion
Sphere	30	(50,100)	0.01
Rastrigin	30	(2.56, 512)	100
Griewank	10	(300,600)	0.05
Griewank	30	(300,600)	0.05
Rosenbrock	30	(15,30)	100
f6	2	(50,100)	0.00001

TABLE III
COMPARISON OF ALGORITHMS ON STANDARDIZED PERFORMANCE AT 1000 ITERATIONS

Type	Standardized Perf.	Perf. S. D.
Canonical	-0.4825	0.0179
FIPS	0.4825	0.8230

TABLE IV
COMPARISON OF ALGORITHMS ON PROPORTION MEETING CRITERIA WITHIN 3000 ITERATIONS

Type	Proportion
Canonical	0.6838
FIPS	0.2839

TABLE V
STANDARDIZED PERFORMANCE AT 1000 ITERATIONS FOR CANONICAL AND FIPS ALGORITHM BY MEAN DEGREE

Type	k=3	k=5	k=10
Canonical	-0.46200	-0.48487	-0.48752
(S.D.)	(0.68362)	(0.70639)	(0.67500)
FIPS	-0.50498	-0.35777	1.09356
(S.D.)	(0.77697)	(0.61482)	(0.01314)

TABLE VI
PROPORTIONS MEETING CRITERIA FOR CANONICAL AND FIPS ALGORITHM BY MEAN DEGREE

Type	k=3	k=5	k=10
Canonical	0.68362	0.70639	0.67500
FIPS	0.77697	0.61482	0.01314

TABLE VII
COMPARISON OF GOOD AND BAD TOPOLOGIES, CLASSIFIED BY PROPORTION MEETING CRITERIA

Type	k=3	k=5	k=10
Canonical			
Good	231	311	801
Bad	0	0	0
FIPS			
Good	209	296	0
Bad	22	15	801

attained by finding the peak of a locally optimal hill. Thus, a second measure was used. The performance criteria were taken from common usage in the literature. These criteria are considered to mean that the algorithm has found the region containing the global optimum; in this experiment, we noted whether the trial successfully met the criterion within 3000 iterations.

Each trial of the algorithm was run for at least 1000 iterations. If the criterion had been met, it stopped there; otherwise, it ran until either 3000 iterations were done or the criterion was met.

Five functions were tested: Sphere, Rastrigin, Griewank, Rosenbrock, and Schaffer's f6 functions. All except f6 were run in 30 dimensions. Schaffer's f6 is a two-dimensional problem. The Griewank function was also run in ten dimensions, as that has been seen to be more difficult than 30. Since the five functions are standard, we have not described these here [5].

All function optima lie at the origin. Therefore, in order to make the test difficult, particles were initialized in a range that did not include the origin. Again, asymmetric initializations were ones widely reported in the literature (e.g., [9]). Criteria and initialization ranges are shown in Table II. Both the algorithms were coded in C, under Linux, using gcc. All of the 1343 topological conditions were tested 40 times on each of the six functions.

VI. RESULTS

Table III shows a comparison of both the algorithms on standardized performance at 1000 iterations. As can be seen, in all the 1343 graphs, the canonical particle swarm performed much better than FIPS. Wherever the canonical version was— on average, nearly half a standard deviation below the mean (remember we are minimizing)—the FIPS version was the same distance above it.

Table IV shows a comparison of both the algorithms on proportion meeting the criteria within 3000 iterations. More than two thirds of the canonical trials met the criteria, while less than one third of the FIPS versions met the criteria.

The means of the two measures, aggregated per graph, correlated very highly, $r(2684) = -0.93, p < 0.0001$. The negative correlation means that topologies that performed well on performance at 1000 iterations—in other words, that had a low value—tended to meet the criteria a high proportion of the time, a high value.

Table V presents a bit of a surprise. It turns out that the FIPS version performed very badly after 1000 iterations under $k = 10$ conditions but performed better than the canonical version under $k = 3$ conditions. The canonical algorithm performed evenly across all graph degrees.

Table VI shows the same pattern when we examine the proportion meeting the criteria. The FIPS with $k = 3$ met the criteria more often than the best canonical version but performed miserably with $k = 10$. Thus, it appears that the FIPS is more susceptible to alterations in the topology, but with a good topology it can outperform the canonical version. Topologies were categorized by whether they met the criteria on half the trials or not.

As can be seen in Table VII, the mean degree of the graph does not explain everything. All the canonical population structures met the criteria at least half the time regardless of k, and none of the FIPS topologies met the criteria when $k = 10$. But it is apparent that the FIPS was not perfect even when the average degree was lower.

Table VIII compares the good and bad FIPS populations when $k = 3$. It is seen that the better graphs had higher clustering of nodes, which means that neighbors tended to connect to neighbors, and also higher variability of clustering. This follows since k is constant throughout these graphs.

Fig. 3 shows the best three topologies for the two algorithms (canonical and FIPS) for each of the two dependent variables. The top nine canonical versions with the highest proportion meeting criteria all had $k = 5$, while seven of the ten best by the standardized performance measure had $k = 10$. Thus, while the large neighborhoods led to quick convergence, they did not lead to convergence on global optima. The best FIPS populations, on the other hand, all had $k = 3$, though the best three by each standard were different from one another.

Table IX shows the three FIPS topologies that were the best on the performance and the proportion measures. It can be seen that the graphs that performed well at 1000 iterations were more highly clustered than the ones that found the global optimum and had much more variance in degree k. The mean distance from one node to another was higher for

TABLE VIII
COMPARISON OF GOOD AND BAD FIPS GRAPHS BY INDEPENDENT VARIABLES

Bad FIPS Graphs		
Variable	**Value**	**S.D.**
Proportion	0.2768939	0.0738529
C	0.0995455	0.3221982
CSD	0.1180336	0.3820388
D	2.5712923	0.3834764
DSD	0.4665250	0.0976014
k	3.0000000	0.0
kSD	1.9016057	0.3184722
Good FIPS Graphs		
Proportion	0.8296053	0.0577391
C	0.6306425	0.3943410
CSD	0.3988160	0.3832100
D	3.0765552	0.5289256
DSD	0.3604364	0.2380061
k	3.0000000	0.0
kSD	0.5296363	0.7565627

C = clustering, D = mean distance between nodes, k = mean degree, and SD = standard deviation.

TABLE IX
STATISTICS FOR THE THREE BEST GRAPHS BY EACH MEASURE

Standardized Performance		
Variables	**Value**	**S.D.**
C	0.989	0
CSD	0.080	0
D	2.530	0.096
DSD	0.445	0.026
k	3.000	0
kSD	2.026	0
Proportion Meeting Criteria		
C	0.525	0.238
CSD	0.501	0.254
D	3.326	0.448
DSD	0.547	0.062
k	3.000	0
kSD	0.108	0.187

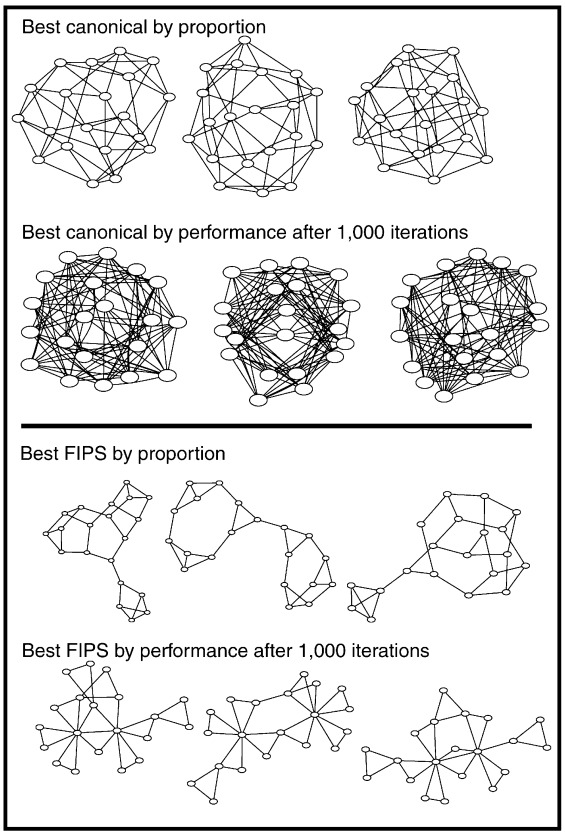

Fig. 3. Three best topologies for each algorithm, classified by performance at 1000 iterations or proportion meeting criteria.

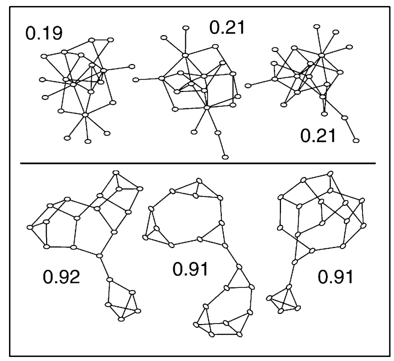

Fig. 4. (Top) Worst and (bottom) best FIPS topologies with $k = 3$ and the proportion of populations containing at least one member meeting criteria.

the ones that found the global optimum. Taken together, these statistics suggest (though n is too small to allow inference) that the inhibition of communication might contribute to the discovery of global optima in the FIPS algorithm, while promotion of communication speeds convergence within optimal regions.

Fig. 4 shows the three worst and the three best FIPS topologies, defined by the proportion meeting the criteria, where $k = 3$. The poorly performing sociograms show numerous isolated individuals, with only one neighbor, and sections of the population that are very highly interconnected.

The high-performing topologies show the effect of clustering described above. These best populations are typified by highly connected subgroups of individuals, which are connected to one another by a small number of links, usually one. The information discovered by a member of the clique is elaborated within the highly interconnected subpopulation, and various cliques operate in parallel, exchanging information through "gateway" members.

The best canonical sociometry met the criteria 80.41667% of the time. Although the FIPS resulted in many failures, 173 different FIPS topologies met the criteria more than the best canonical topology. Moreover, while the best-performing canonical version, as measured by the aggregated standardized performance at 1000 iterations, was 0.5249531 standard deviations below the mean, 85 different FIPS sociometries performed better than that.

VII. CONCLUSION

The particle swarm algorithm works through influence among particles. The previous versions of the algorithm have assumed that particles should receive the best information available to them and that they should be influenced by the best point found by the best member of their neighborhoods. But this does not seem to be true.

The neighborhood provides two kinds of information to the target particle. First, and most obviously, the locations where neighbors have found relatively good solutions represent good places in the search space. The pattern of variables that has given a particle its very best solution is probably, on a regular function space, indicative of at least a local optimum. It is quite likely that the best points found by a target's neighbors are from different regions; however, this would seem to suggest that the particle would be less confused if it were influenced only by a single neighbor at a time, as in the canonical version.

The second kind of information conveyed by the neighborhood is equally important. The distances between the particles indicate how much consensus has been attained and determines the size of the particle's steps through the search space. In the canonical version, the velocity is adjusted by an amount that is determined by the distance between two points: An individual's previous best and the neighborhood's previous best. In the FIPS, the velocity is adjusted by an amount that is a kind of average difference between each neighbor's previous best and the target particle's current position. In either case, the neighborhood consensus tells the particle how big its steps through the search space should be.

The change in optimal stepsize has been described in the literature as a shift from exploration to exploitation, i.e., from search for an optimal region to search for an optimal point within the region. This dichotomy is useful on search spaces where local optima are general and well distinguished from one another but may not apply on corrugated or otherwise irregular search spaces. The particle swarm tends to decrease stepsize as an optimal point is reached because of the consensus among neighborhood members, but if a good point that is remote in the search space is reported, a particle's trajectory can return to its exploratory mode. Moreover, in this paradigm the two modes of search are considered as continuous, matters of degree rather than dichotomy, simply scaling the length of steps taken.

In the FIPS, the target individual's previous best position is not included in the computations, but only the discrepancy between the neighbors' best and their current positions. Thus, the FIPS is a stronger social influence model than the canonical version. Note that there is no social-psychological principle that says that an individual's experience contributes half to his or her learning. In fact, studies such as Mullen's [8] suggest that an individual's experience tends to be overwhelmed by social influence.

It is not possible to say, from these data, which version of the particle swarm is better. It seems that both performed well. The FIPS provided both the worst and the best performances in this data set and the performance was highly contingent on the population topology. While the FIPS had many topologies that outperformed the best canonical versions, by both measures, the preponderance of FIPS trials were abysmal. To compare, it will be necessary to find the topologies that are very best suited to each algorithm.

References

[1] M. Clerc and J. Kennedy, "The particle swarm: Explosion, stability, and convergence in a multi-dimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, Feb. 2002.

[2] M. S. Granovetter, "The strength of weak ties," *Amer. J. Sociol.*, vol. 78, pp. 1360–1380, 1973.

[3] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Human Sci.*, Nagoya, Japan, 1995, pp. 39–43.

[4] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, Australia, 1995, vol. 4, pp. 1942–1948.

[5] J. Kennedy, R. C. Eberhart, and Y. Shi, *Swarm Intelligence*. San Francisco, CA: Morgan Kaufmann/Academic, 2001.

[6] J. Kennedy and R. Mendes, "Population structure and particle swarm performance," presented at the 2002 World Congress on Computational Intelligence, Honolulu, HI, May 12–17, 2002.

[7] B. Latané, "The psychology of social impact," *Amer. Psychol.*, vol. 36, pp. 343–356, 1981.

[8] B. Mullen, "Operationalizing the effect of the group on the individual: A self-attention perspective," *J. Exp. Social Psychol.*, vol. 19, pp. 295–322, 1983.

[9] Y. Shi and R. Eberhart, "Experimental study of particle swarm optimization," presented at the 4th World Multiconf. Systemics, Cybernetics and Informatics (SCI2000), Orlando, FL, 2000.

Using Genetic Algorithms to Estimate Confidence Intervals for Missing Spatial Data

Neil H. W. Eklund, *Member, IEEE*

***Abstract*—Gas turbine blades, which come in many shapes and sizes, must meet strict engineering specifications. The current manual blade measurement system is slow and labor intensive. As part of the development of an optical measurement system, an approach for characterizing missing data was required. A novel technique for conditional spatial simulation using genetic algorithms (GAs) was developed. The problem is encoded using the "random key genetic algorithm" (RKGA) approach. The RKGA allows the use of a sampling distribution for missing measurements that can accommodate values uncharacteristic of the area surrounding the missing data, while still allowing realizations of the missing data with reasonable directional semivariance characteristics to be developed. A unique optimization approach was used, consisting of a crossover-only GA, followed by a hill-climbing phase. Each phase addresses different parts of the problem (the low and high special frequencies, respectively). This spatial simulation technique can be used to characterize regions of missing data in regularly sampled measurements. The proposed technique is much faster than simulated annealing, the current state of the art in spatial simulation. An application of this technique to determining confidence intervals for missing data in optical measurements of gas turbines is described.**

***Index Terms*—Gas turbine blades, manufacturing, optical measurement, spatial simulated annealing (SSA).**

I. Introduction

Gas turbine blades, which come in many sizes and shapes (even within one turbine), must be measured to ensure that they meet strict engineering specifications, both when they are originally manufactured and when the turbine is serviced. Using mechanical methods to measure blades to the required accuracy is very slow; it would be much faster to use optical methods, which map the surface of the blade using images of laser light directed at a lattice of points on the blade [18]. However, missing data is a problem for optical measurement methods [17]: while measurements can be made where the surface of the blade is diffused, where the surface of the blade is highly specular an image is not returned to the detector, resulting in patches of missing measurements.

The surface of a turbine blade is quite smooth, and so uncertainty resulting from interpolation is acceptable over a small region of missing data. However, for larger regions of missing data, the potential difference between the estimated surface and the actual surface can be so great as to require the rejection of the measurement. The goal

Manuscript received March 5, 2003; revised December 20, 2005; accepted December 28, 2005. This paper was recommended by Guest Editor S. W. Kercel.

The author is with the Industrial Artificial Intelligence Laboratory, GE Global Research Center, Niskayuna, NY 12309 USA (e-mail: eklund@research.ge.com).

Digital Object Identifier 10.1109/TSMCC.2006.875407