Computational Intelligence in Radio Astronomy: Using Computational Intelligence Techniques to Tune Geodesy Models

Daniel Angus[1,⋆] and Adam Deller[2]

[1] The University of Queensland, Queensland Brain Institute
d.angus@uq.edu.au
[2] Swinburne University of Technology, Centre for Astrophysics and Supercomputing
adeller@astro.swin.edu.au

Abstract. In this paper a number of popular Computational Intelligence (CI) algorithms are used to tune Geodesy models, a radio astronomy problem. Several single and multiple objective variations of the Geodesy problem are examined with good results obtained using state-of-the-art CI algorithms. These novel applications are used to develop insights into methods for applying CI algorithms to unknown problem domains and to provide interesting solutions to the Geodesy models used.

1 Introduction

Astronomy, one of the oldest of the sciences has many important problems with far reaching implications. Radio astronomy, which involves the observation of radiation usually greater than one millimetre is important in observing particular celestial objects, such as pulsars. To observe such phenomena, radio astronomers may make use of large scale, spatially separated radio antenna arrays, a technique known as Very Long Baseline Interferometry (VLBI). The use of such equipment comes with certain restrictions though due to factors such as random atmospheric noise and fundamental equipment limitations. Because of this, methods to limit the contribution of noise and errors and thus enhance the quality of observations are important.

One potential source of error in observation relates to the precise location of the actual observational antenna stations. Due to tectonic plate movement and other forces the positions of these antennae are in constant motion. Geodesy is a field which deals directly with measuring and studying the earth's geometry thus it is a useful to apply geodesy techniques to the problem of accurately predicting the positions of antennae.

Computational models that describe these position errors are available, however due to the complexity of the system under investigation it is not always known what parameters should be used in such models. Previously, such models

⋆ This authors' research was partly sponsored by Swinburne University of Technology, Complex Intelligent Systems Lab.

X. Li et al. (Eds.): SEAL 2008, LNCS 5361, pp. 615–624, 2008.

relied on simple deterministic solvers checking through vast numbers of parameter combinations to determine the best parameterisation, however it is obvious that such solutions can be computationally expensive.

Computational Intelligence (CI) is a broad term used to describe many different algorithms, classification systems and computational models usually based on observed physical phenomena. Examples of CI techniques include Evolutionary Algorithms, Swarm Intelligence and Artificial Neural Networks. These problem solvers can often be fast and computationally inexpensive as they tend to adjust their search behaviour dynamically during run-time to better exploit promising areas of the search space. Such techniques have been shown to be effective on complex, multi-modal search spaces typical of that of the geodesy problem described above.

In this paper a range of popular CI techniques are used to optimise a geodesy model. Various techniques aimed at improving the efficacy of the results are explored and the final results are shown to be competitive with an alternative, computationally expensive, deterministic solver. The paper begins by explaining the problem in detail in Sec. 2. The algorithms and their specific application strategies are explained in Sec. 3. Results are provided in Sec. 4, and concluding remarks offered in Sec. 5.

2 Problem Details

2.1 Problem Description

The radio astronomy problem described in this section is taken from the field of Very Long Baseline Interferometry (VLBI). The solution of this problem is significant to radio astronomers since an accurate geometric model is essential to correctly align sampled data sequences from different antennas in time for any interferometry experiment. Errors in antenna positions result in a time-variable offset between the data streams, which reduces image fidelity and introduces systematic errors to source position estimates.

Radio interferometry involves the temporal alignment of signals from different radio telescopes (by means of electronic delays) and correlation of the signals for the purpose of determining information about the spatial frequencies of the radio sky within the telescope's field of view. The collected data is later Fourier transformed to create an image of the radio sky. This data stream alignment requires each telescope's geocentric (geographical) position to be precisely determined, and is made more difficult due to the addition of propagation delays through the Earth's atmosphere, time-stamping errors, the source position and structure, and Earth orientation.

The residual (unmodelled) delay between a pair of antennas (a baseline) can be estimated by fitting the correlated signal phase as a function of frequency, a process known as *fringe-fitting*. For typical astronomical observations, the residual delay (usually resulting from unmodelled atmospheric contributions and timestamp errors) is simply an error that is removed, and its source is immaterial. However, it is possible to use the residual delays to determine errors in antenna

$$\delta_{B-A} = ((\mathbf{B}' \cdot \hat{\mathbf{S}}' - \mathbf{B} \cdot \hat{\mathbf{S}}) - (\mathbf{A}' \cdot \hat{\mathbf{S}}' - \mathbf{A} \cdot \hat{\mathbf{S}}))/c + (C_B - C_A) + (A_A - A_B) + N_{BA} \quad (1)$$

Where:

$\mathbf{B}$, $\mathbf{A}$	: Geocentric antenna vectors (user defined variables)
$\hat{\mathbf{S}}$	: Unit vector toward the source (user defined)
$\mathbf{B}'$, $\mathbf{A}'$	: Geocentric antenna vectors (observed values)
$\hat{\mathbf{S}}'$	: Unit vector toward the source (observed value)
c	: Speed of light
C	: Time-stamp error at an antenna
A	: Unmodelled delay due to atmospheric propagation
N_{BA}	: Error introduced by finite signal/noise.

positions, a procedure known as *geodesy*. Geodesy is a challenging problem, since attaining millimetre accuracy requires measurement of relative delays on the order of picoseconds, and the removal of all other contaminating effects on the residual delay. The value of residual delay δ from antenna **B** to antenna **A** obtained from the fringe-fit solutions can be written as Equ. 1.

Of these values, N_{BA} and C have no dependence on antenna or source position. N_{BA} is randomly distributed and C is a smoothly varying function of time with modern frequency standards, leaving the geometric ($\mathbf{A}$,$\mathbf{B}$,$\hat{\mathbf{S}}$) and propagation terms (A). In a dedicated geodesic experiment, weather information is monitored at all antennas and used to correct *a priori* atmospheric propagation models, leaving antenna position errors as the dominant contribution to residual delay. However, even if atmospheric and source position errors cannot be completely subtracted from the residual delays, an estimate of antenna position errors can still be made by assuming the atmospheric and source error delay contributions to be uncorrelated with baseline error delay contributions. An example of an observed residual delay for a single baseline tracked over a 24 hour period is presented in Fig. 1[1].

2.2 Objective Function

From an optimisation perspective the interest is in minimising the magnitude of the residual delay per baseline (antenna pair). This is achieved in this case by selecting appropriate values for the vectors $\mathbf{A}$,$\mathbf{B}$ and $\hat{\mathbf{S}}$. Like any black box optimisation problem though these vectors are unknown and as such trial and error must be used to select their values. These vectors translate into 14 configurable (scalar) parameters per baseline. However, since one antenna is used as a common reference antenna for all antennae in the system the actual number of configurable parameters per antenna is 7. These values represent the x, y, and z positions, two clock offsets (one per observed frequency band), a clock rate (electronic delay) and clock acceleration (drift in timestamp error). The numbers are

[1] A far more rigorous description can be found in [1].

Table 1. Search space range per individual problem dimension for all datasets

Parameter	Lower bound	Upper bound	Range
1,2,3	-1.0	1.0	2.0
4,5	-5.0E-9	5.0E-9	1.0E-8
6,7	-5.0E-8	5.0E-8	1.0E-7

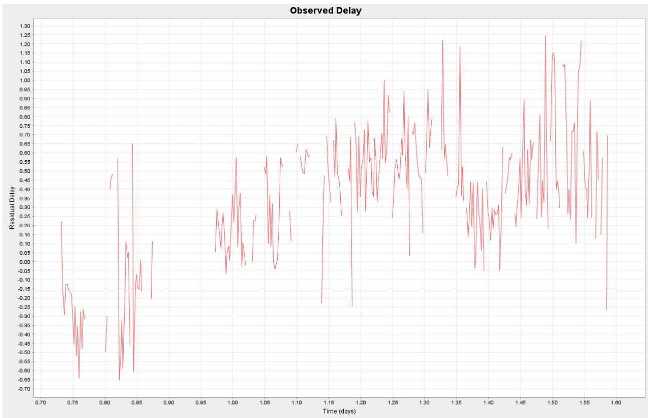

Fig. 1. The observed residual delay for one observing band of an antenna pair (Parkes and ATCA) tracked over a 24 hour period taken from data set STA-131AV

real numbers limited within the range given in Tab. 1 and are valid within this entire range.

The objective value for each baseline is calculated by measuring the difference between each predicted data point (derived from the 14 parameters selected per baseline) and its corresponding observed (real) data point. These differences are combined in an RMS fashion into a single composite objective value. An optimal parameter selection would allow the predicted and observed residual delays to overlap each other perfectly such that if one was subtracted from the other they would essentially cancel each other out, however due to the random noise term N_{BA} this is impossible, and a solution with minimal difference is deemed optimal.

2.3 Problem Data

The problem data used was gathered from five observation points, Parkes, ATCA, Mopra, Hobart and Ceduna. For all data sets Parkes is used as the reference antenna as it is the most sensitive, thus minimising the contribution of the noise term N_{BA}. The data sets STA-131AU and STA-131AV were based on observations of two different sets of radio sources and are available upon request from the Australia Telescope National Facility[2]. Use of these datasets gives rise to four separate baselines that all use Parkes as a common reference.

[2] Contact: `vlbi@atnf.csiro.au`

2.4 Multiple and Single Objective Variations

There are a variety of ways to attempt such an optimisation problem. If all four baselines are optimised in parallel, the problem can be phrased as a multiple or single objective optimisation problem. In the multiple objective optimisation (MOO) variant all input parameters (35 in total[3]) are varied at once, while the objective values for each baseline are treated as independent values to be minimised, denoted as 35D-MO. The single objective optimisation (SOO) variant would require the same number of input parameters, however the objective values can be combined in an RMS fashion to obtain a single objective value, denoted as 35D-SO. Alternatively the seven parameters for the reference antenna can be fixed at zero, resulting in the separation, or decoupling, of the baselines into four separate single objective, 7 dimension, CFO problems, denoted as 7D-SO.

3 Algorithm Description

3.1 Algorithms Used

Multiple objective algorithms used in this study include the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [2] and a Population-based Ant Colony Optimisation algorithm (PACO-MOFO) [3]. The NSGA-II algorithm is a state-of-the-art MOO algorithm which extends many of the basic functions of the canonical Genetic Algorithm. It uses a dominance ranking procedure and crowding distance comparison to determine its population composition, and uses traditional Genetic Algorithm procedures such as crossover and mutation in new solution construction. The PACO-MOFO algorithm is a recently developed MOO algorithm which uses a stepwise solution construction procedure, like traditional Ant Colony Optimisation techniques, to create new solutions. The PACO-MOFO algorithm uses dominance ranking to score solutions and a crowding replacement operation is used to add new solutions to the existing population. These algorithms have been selected due to their success in past studies when applied to complex multiple objective function optimisation problems.

Single objective algorithms include Particle Swarm Optimisation (PSO) [4], Crowding Population-based Ant Colony Optimisation (CPACO) [5], Deterministic Crowding Genetic Algorithm (DCGA) [6,7] and Differential Evolution (DE) [8]. PSO is a optimisation technique inspired by the flocking of birds, it uses a vector-based approach to direct a population of solutions through the search space. The solutions are evaluated periodically to determine whether new interesting optima have been located. CPACO is another Population-based Ant Colony Optimisation algorithm which uses a stepwise solution construction mechanism to construct solutions one dimension at a time. Good solutions are stored in a population and each solution manipulates an artificial ‘pheromone’

[3] 7 parameters are specified per antenna (including the reference). Since the same parameters for the reference antenna are used by each baseline this results in 35 dimensions for a 5 antennae problem.

which is used to bias the construction of new solutions. DCGA is a Genetic Algorithm which uses a real-value encoding and a deterministic crowding replacement operation. Two solutions are created at a time from two parent solutions using crossover and mutation operators. The solutions are then compared directly to each of these parents and if either or both of the new solutions are better than their closest matching parent, the parent is replaced. The DE algorithm is similar to PSO as it is also uses a vector based approach to search the given search space. In its simplest form DE selects three population members at random, the difference between two is calculated and used to perturb the third. If the new perturbed solution is better than a selected existing population member that population member is replaced.

3.2 Methodology Employed

Of particular interest in this study is the development of methods to increase the search efficacy of the previously described CI techniques. Given that very little is known about the search space composition and dynamics the development of general purpose solution solving methods is of interest. The particular methodologies that are tested include:

Search Space Size Reduction. It is intuitive that a reduction in the size of the search space will lead to a much more thorough search given a limited number of function evaluations. However as it is not known where the actual global optimal solution is located the reduction of the search space may actually exclude this global optimal solution thus possibly decrease the efficacy of a global search strategy. The search space size is reduced using a first-order solution estimation technique which performs a rough fit of the data and centers the search space around this solution.

Search Space Relocation. Given that a reduction in the size of the search space may lead to the exclusion of a global optimal solution, a possible alternative is to allow the search space to periodically relocate. It is thought that such a relocation may, for some period of time, allow the inclusion of more promising areas of the search space thus increasing the overall search efficacy. To achieve the relocation each algorithm is allowed m function evaluations after which the best solution is used as an origin where the search space is relocated around. The actual size of the search space does not change, simply the minimum and maximum boundaries are shifted. After repositioning the search space boundaries the algorithm is reinitialised and allowed to run again. This process is repeated p times so that overall the algorithm executes $m \times p$ function evaluations. The best solution found during the entire run is reported rather than the last best solution found.

Problem Variations. In the 35D-MO, 35D-SO configurations the problem is treated as a $7n$ dimension problem, where n is the number of antennae, inclusive of the reference antenna. Alternatively, if the values for the reference antenna are fixed at zero this allows the problem to be decomposed into several smaller (less dimensions) problems since the effect of a variable reference is removed. Removal of the reference antenna's variability is a valid approach since baseline measurements are always relative to this reference antenna anyway, as such it means that all baseline variation is occurring at an individual antenna rather than being split between the reference and selected antennae. By fixing the reference antenna the problem can be treated as four independent 7 dimension problems. The complexity of evaluating the objective value of these variations is dependent on the number of input variables. As a result the 35D-MO, 35D-SO and 7D-SO (evaluated for all four baselines) all require a comparable amount of computation to evaluate solutions.

4 Results and Analysis

4.1 Experiment Setup

The 35D-MO and 35D-SO configurations were tested first, using the default variable ranges defined in Tab. 1. The 35D-SO and 7D-SO configurations were then tested, still using the default variable ranges, to determine the effect of separating the problem into multiple sub-problems[4]. A reduction of the search space size was then tested using only the 7D-SO problem variant. Finally a reduction of the search space size with dynamic search space relocation was performed using the 7D-SO problem variant.

All algorithms were allowed a total of 100,000 function evaluations per trial, and all trials were repeated 100 times using different random seeds. Non-parametric statistics were used to validate significant differences between reported results, specifically the Mann-Whitney rank sum test was used with a confidence level of 95%. The results reported are the mean (with standard deviation) and those highlighted in bold were the best values found. Algorithm parameter configurations were set using the suggested default values in the original references, and with population sizes of 100.

4.2 Experiment Results

Given that the objective space of the 35D-MO variant contains four dimensions, it was difficult to visualise the results obtained by the NSGA-II and PACOMOFO algorithms using traditional visualisation methods such as summary attainment surface comparison. Instead the objective values were combined for each individual solution and the best aggregate value is reported in Tab. 3 to provide values that are directly comparable to the results obtained by the algorithms used in the 35D-SO and 7D-SO configurations reported in Tab. 2.

[4] For the 7D-SO model the algorithms use four independent 7D objective functions, however, the value reported is obtained by recombining the best solutions' input parameters found for each sub-problem and evaluating this combination using the 35D objective function.

Table 2. Results obtained by the Random, PSO, DE, CPACO and DCGA algorithms, on the 35D-SO and 7D-SO variants using the default search space size

Problem	Random	PSO	DE	CPACO	DCGA
AU(35D-SO)	12.60 (1.44)	5.79 (1.72)	**0.76 (0.07)**	1.19 (0.16)	2.41 (0.42)
AU(7D-SO)	-	1.16 (0.23)	**0.54 (0.00)**	0.60 (0.04)	0.59 (0.01)
AV(35D-SO)	12.66 (1.62)	5.71 (1.47)	**0.68 (0.07)**	1.17 (0.21)	2.30 (0.53)
AV(7D-SO)	-	1.07 (0.30)	**0.48 (0.00)**	0.53 (0.02)	0.51 (0.01)

Table 3. Results obtained by the NSGA-II and PACO-MOFO algorithms on the 35D-MO variant

Problem (metric)	NSGA-II	PACO-MOFO
AU (combined score)	39.84 (14.02)	**8.36 (1.81)**
AV (combined score)	43.09 (18.73)	**8.89 (2.25)**

Table 4. Results obtained by the Random, PSO, DE, CPACO and DCGA algorithms on the 35D-SO and 7D-SO variants using heuristic search space reduction

Problem	Random	PSO	DE	CPACO	DCGA
AU(35D-SO)	6.13 (1.05)	0.65 (0.04)	**0.55 (0.00)**	0.68 (0.05)	1.30 (0.35)
AU(7D-SO)	-	0.59 (0.00)	0.59 (0.00)	0.59 (0.00)	0.59 (0.00)
AV(35D-SO)	6.27 (1.14)	0.57 (0.04)	**0.50 (0.00)**	0.62 (0.06)	1.22 (0.36)
AV(7D-SO)	-	0.51 (0.00)	0.51 (0.00)	0.51 (0.00)	0.51 (0.00)

Table 5. Results obtained by the DE and CPACO algorithms on the 7D-SO variant using heuristic search space reduction and dynamic search space re-allocation

Problem	DE	CPACO
AU(35D-SO)	**0.60 (0.02)**	0.70 (0.13)
AU(7D-SO)	**0.52 (0.00)**	**0.52 (0.00)**
AV(35D-SO)	**0.56 (0.02)**	0.61 (0.07)
AV(7D-SO)	**0.48 (0.00)**	**0.48 (0.00)**

Table 6. Positional errors obtained for the STA-131AV dataset using an alternative (but more complex) geodesy model with a deterministic solver. Positional errors obtained for the STA-131AV dataset using DE and CPACO on the 7D-SO problem configuration with the heuristic search space reduction and dynamic relocation technique. Magnitude of difference between the positional errors.

	Deterministic Solver			DE / CPACO			|Difference|		
Baseline	x (m)	y (m)	z (m)	x (m)	y (m)	z (m)	x (m)	y (m)	z (m)
ATCA	-0.08	0.13	0.31	-0.16	0.14	0.22	0.08	0.01	0.09
Mopra	-0.08	0.08	0.31	-0.15	0.16	0.13	0.07	0.08	0.18
Hobart	0.00	0.18	-0.07	-0.15	0.36	-0.31	0.15	0.18	0.24
Ceduna	-1.72	1.65	-0.89	-1.54	1.47	-0.58	0.18	0.18	0.31

4.3 Experiment Analysis

The results presented in Tab. 2 indicate that when using the default search space range, DE is the best performing algorithm. These results also indicate that when using the default range all algorithms perform better when the reference antenna is fixed at zero and the problem is decoupled into 4 smaller sub-problems (7D-SO). For the Multiple Objective variation (35D-MO) the results (Tab. 3) seem to be quite poor, most likely because these algorithms focus attention across a larger area of the search space (to locate multiple Pareto optimal solutions) when compared to the single objective algorithms. Considering that the problem objectives are easily decomposed into separate sub-problems by fixing the reference antenna at zero, as in 7D-SO, it is unlikely for 35D-MO to be competitive.

When a heuristic search space reduction technique is used, the results for both the 35D-SO and 7D-SO variations for all but the DE algorithm are improved (Tab. 4). The results for DE are interesting because the 35D-SO results are improved using the heuristic search space reduction, however the 7D-SO results re made worse through the reduction of the search space size. This result indicates that the heuristic may exclude optimal areas of the search space. Such a result in the case of DE is bad because this algorithm is able to find good solutions on the larger search space, however in the case of the other algorithms tested such an exclusion is not bad considering these algorithms don't actually find the optimal solution anyway. A similar result was found in a previous CI algorithm study[9]. The optimal search space exclusion is reversed when the search space relocation technique is applied (Tab. 5), and in this case the best results for all experiments performed are obtained by the DE and CPACO algorithms. The agreement between the results obtained by both the DE and CPACO algorithms on the 7D-SO problems using search space reduction and relocation suggests that a global optimal solution has been located.

A deterministic technique operating on a more complex geodesy model was used previously to compute the positional errors for the STA-131AV dataset. This technique used exact measurements of external factors such as the atmospheric noise in its calculations, resulting in a much higher fidelity solution. The results for the positional errors are reported for all four antennas (using Parkes as the reference) and as a comparison the best result obtained with the DE and CPACO algorithms using the 7D-SO variation with heuristic search space size reduction and dynamic relocation (Tab. 6).

Given that the estimated error in the deterministic techniques measurements are approximately ±10cm per dimension, the DE/CPACO results obtained for both ATCA and Mopra are deemed comparable. For the Hobart and Ceduna baselines the DE/CPACO positional errors are larger than the allowed ±10cm, which is postulated to be due to unaccounted systemic errors in the geodesy model rather than the optimisation algorithm not obtaining a good solution. These systemic errors could manifest themselves in a variety of ways since the deterministic model uses cleaner radio data, models the atmosphere better thus reducing introduced noise and uses multiple widely spaced frequency bands which results in an increase in sensitivity.

5 Conclusion

This study has shown the usefulness of CI algorithms in yielding quality results on a previously unstudied real-world optimisation problem. The results obtained by the DE and CPACO algorithms were in some cases shown to be within error bounds of previously determined values using a more accurate geodesy model. Perhaps the most informative result obtained related to the effect of heuristics on the accuracy of the result obtained. For PSO, DCGA and CPACO a reduction in the search space size advantaged their search performance, while for DE such a reduction hindered its search performance. Aside from the quantitative results presented, the actual exercise of applying CI algorithms to the problem was able to provide the Astrophysics domain expert with new insight into the underlying problem dynamics. Such insights are hard to quantify and thus are not reported, however this result indicates an important reason as to why one might consider the application of CI algorithms to their problem.

References

1. Fomalont, E.B.: Astrometry and Geodesy. In: Taylor, G.B., Carilli, C.L., Perley, R.A. (eds.) Synthesis Imaging in Radio Astronomy II. Astronomical Society of the Pacific Conference Series, vol. 180, p. 463 (1999)
2. Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In: Deb, K., Rudolph, G., Lutton, E., Merelo, J.J., Schoenauer, M., Schwefel, H.-P., Yao, X. (eds.) PPSN 2000. LNCS, vol. 1917, pp. 849–858. Springer, Heidelberg (2000)
3. Angus, D.: Population-based ant colony optimisation for multi-objective function optimisation. In: Randall, M., Abbass, H.A., Wiles, J. (eds.) ACAL 2007. LNCS, vol. 4828, pp. 232–244. Springer, Heidelberg (2007)
4. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
5. Angus, D.: Niching for Population-based Ant Colony Optimization. In: 2nd International IEEE Conference on e-Science and Grid Computing, Workshop on Biologically-inspired Optimisation Methods for Parallel and Distributed Architectures: Algorithms, Systems and Applications (2006)
6. Mahfoud, S.W.: Crowding and preselection revisited. In: Männer, R., Manderick, B. (eds.) Parallel Problem Solving from Nature 2 (PPSN2), pp. 27–36. North-Holland, Amsterdam (1992)
7. Mahfoud, S.W.: Niching methods for genetic algorithms. PhD thesis, University of Illinois (1995)
8. Storn, R., Price, K.: Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical Report TR-95-012, ICSI (1995)
9. Montgomery, J.: Alternative representations for the job shop scheduling problem in ant colony optimsation. In: Randall, M., Abbass, H.A., Wiles, J. (eds.) ACAL 2007. LNCS (LNAI), vol. 4828, pp. 1–12. Springer, Heidelberg (2007)