Ecological Optimization Using Harmony Search

ZONG WOO GEEM, JUSTIN C. WILLIAMS
Environmental Planning & Management Program
Johns Hopkins University
Baltimore, Maryland, USA

Abstract: - The music-inspired meta-heuristic algorithm, harmony search, was applied to a natural reserve selection problem for preserving species and their habitats. The problem was formulated as an optimization problem (maximal covering species problem; MCSP) to maximize covered species with minimal efforts. Then, it was solved by an improved harmony search (HS) algorithm which includes problem-specific operations. When applied to real-world problem in the state of Oregon, USA, the harmony search algorithm found better solutions than those of another meta-heuristic algorithm, simulated annealing.

Key-Words: - Harmony search, Optimization, Maximal covering species problem, Evolutionary algorithm, Meta-heuristics, Soft computing

1 Introduction

In modern industrial and urbanized life, conserving ecosystem and its species is very important. In order to do so, quantitative optimization techniques have been developed and utilized for the nature reserve site selection problem.

In the light of the optimization, ReVelle et al. [1] reviewed five classes of the reserve selection problem: (1) species set covering problem (SSCP); (2) maximal covering species problem (MCSP); (3) backup and redundant covering problem (Maximal Multiple-Representation Species Problem; MMRSP); (4) chance constrained covering problem; and (5) expected covering problem. Out of the above-mentioned five classes, the MCSP was especially tackled by various algorithms [2-4].

In this study, we apply a recently-developed harmony search (HS) algorithm to the MCSP and to compare our results with those of another meta-heuristic algorithm, simulated annealing (SA), from the literature [2].

2 Problem Formulation

Of the five classes of reserve selection problems discussed in [1], SSCP and MCSP are especially popular forms. The SSCP is to find the least number of parcels while covering every species.

The mathematical formulation of the SSCP model is as follows:

$$Min \sum_{j \in J} x_j \tag{1}$$

$$s.t. \sum_{j \in M_i} x_j \geq 1, \quad all\ i \in I. \tag{2}$$

where j and J are the index and set of land parcels, respectively; i and I are the index and set of species, respectively; M_i is the set of parcels j that include species i; and x_j is a binary variable for parcel selection (it has 1 if parcel is selected, and has 0 otherwise).

The MCSP is to find the maximal number of species while limiting the number of selected parcels to P. The mathematical formulation of the MCSP model is as follows:

$$Max \sum_{i \in I} y_i \tag{3}$$

$$s.t. \sum_{j \in M_i} x_j \geq y_i, \quad all\ i \in I, \tag{4}$$

$$\sum_{j \in J} x_j = P. \tag{5}$$

where y_i is a binary variable for species covering (it has 1 if species i is covered, and has 0 otherwise).

ISSN: 1790-5117

ISBN: 978-960-6766-47-3

3 Harmony Search Algorithm

The HS algorithm originally came from the analogy between music improvisation and optimization process [5]. This algorithm has been successfully applied to various discrete optimization problems such as traveling salesperson problem [5], tour routing [6], music composition [7], Sudoku puzzle solving [8], water network design [9], dam operation [10], vehicle routing [11], and structural design [12].

The HS algorithm searches for optimal solution vectors using a novel stochastic derivative as in Equation 6 [13], which is based on the density information of multiple solution vectors, instead of gradient information of single solution vector. In other words, if a certain value frequently appears in multiple vectors, the value has higher chance to be selected.

$$\begin{aligned}\frac{\partial f}{\partial x_i} &= \frac{1}{K_i}\cdot P_{Random} \\ &+ \frac{n(x_i(k))}{HMS}\cdot P_{Memory} \\ &+ \frac{n\big(x_i(k \mp m)\big)}{HMS}\cdot P_{Pitch}\end{aligned} \tag{6}$$

The stochastic derivative in Equation 6 stands for the total probability (random selection probability + memory consideration probability + pitch adjustment probability) to select a certain value $x_i(k)$ from candidate discrete value set $\{x_i(1), x_i(2), ..., x_i(K_i)\}$ or from a harmony memory (HM) $\{x_i^1, x_i^2, ..., x_i^{HMS}\}$ which stores multiple solution vectors. Here, $f(\cdot)$ is an objective function; $n(\cdot)$ is a frequency function which counts the number of a certain value $x_i(k)$ in the HM; k is an index; K_i is the number of candidate values for the variable x_i; m is a neighboring index and has normally 1; and HMS is the number of multiple vectors stored in the HM.

The first term in the right hand side in Equation 6 indicates the probability to randomly select $x_i(k)$ from candidate value set with the probability of P_{Random}; the second term the probability to select $x_i(k)$ from the HM with the probability of P_{Memory}; the third term the probability to select $x_i(k)$ after its neighboring values are selected from the HM, with the probability of P_{Pitch}.

The basic procedure of HS is as follows:

- Step 0. Randomly generate multiple solution vectors as many as HMS (harmony memory size).
- Step 1. Generate a new vector based on the stochastic derivate in Equation 6.
- Step 2. If the new vector is better than the worst one in the HM in terms of an objective function, the new one is included in the HM and the worst one is excluded from the HM.
- Step 3. Repeat Steps 1 and 2 until stopping criteria is satisfied.

For example, consider a minimization problem as follows:

$$Min \quad f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 3)^2 \tag{7}$$

In Step 0, two vectors (3, 2) and (4, 5) are randomly generated with the candidate value set of $\{1, 2, 3, 4, 5\}$. The HM can be as follows:

$$\mathbf{HM} = \left[\begin{array}{cc|c} 3 & 2 & 2 \\ 4 & 5 & 8 \end{array}\right] \tag{8}$$

In Step 1, the value of x_1^{New} can be 4 if it is chosen from the HM {3, 4}; and the value of x_2^{New} can be 3 if 2 is first chosen from the HM {2, 5}, then 2 is pitch-adjusted into a neighboring value 3. The new vector (4, 3) has the objective function value of 4.

In Step 2, the new vector (4, 3) is included in the HM and the worst vector (4, 5) is excluded from the HM because the former has better value than the worst one.

4 Application

The HS model developed for the MCSP is tested with Oregon data, which consists of 426 species and 441 parcels. Figure 1 shows the hex map of the state of Oregon.

For being applied to the MCSP, the structure of the original HS algorithm is modified. Normally pitch adjustment operation is important because it helps the algorithm to locally search. However, because the decision variable x_j has only two candidate values $\{0,1\}$, it (third term in Equation 6) is omitted.

Also, because the value of P (which ranged from 1 to 24) in Equation 5 is relatively small when compared with the number (441) of entire variables, the decision variable x_j should have 1 with the following probability, instead of 50 %, in initial multiple vector generation (Step 0) and random selection operation (first term in Equation 6):

$$x_j \leftarrow \begin{cases} 1 & w.p. \quad P\dfrac{\frac{N_j}{N_t}}{\sum_j \frac{N_j}{N_t}}, \quad j=1,\cdots,J \\ 0 & \quad otherwise \end{cases} \tag{9}$$

where N_j is the number of species that are included in parcel j; and N_t is the number of total species (= 426 in this example) eligible to be covered.

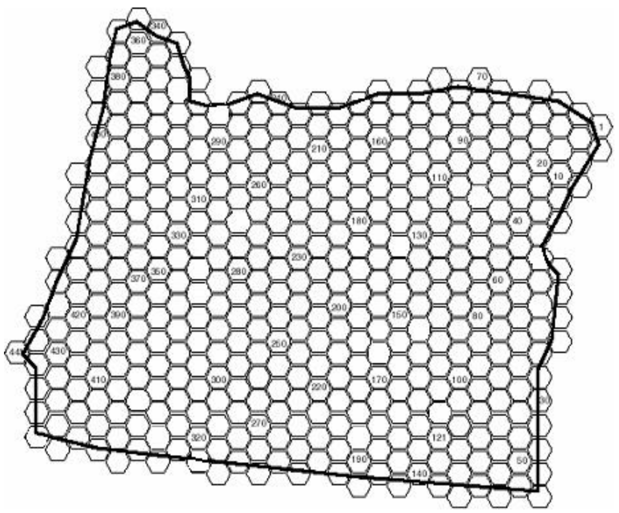

Figure 1. Hex Map of Oregon

Equation 9 allows us to efficiently generate a proper solution vector in sparse selection. In other words, it helps the algorithm to choose only 1 to 24 parcels for a solution vector which has 441 candidate parcels.

However, some vectors have selections (1's) less than P times. For this reason, a coefficient α (it has 1.5 in this study. The numerical results with 1.5 were better than those with 1.0) is added to Equation 9 as follows:

$$x_j \leftarrow \begin{cases} 1 & w.p. \quad \alpha\left(P\dfrac{\frac{N_j}{N_t}}{\sum_j \frac{N_j}{N_t}}\right), \quad j=1,\cdots,J \\ 0 & \quad otherwise \end{cases} \tag{10}$$

If the number of selected parcels reaches P, HS stops considering additional selection using Equation 10 and goes to Step 2 in the HS procedure. In order not to lose parcels which have more species than others, each decision variable is considered in descending order of selection rate in Equation 10, instead of sequential order of j.

In order to start with better initial vectors in Step 0, the HM is filled with chosen multiple vectors, as many as HMS, after (n × HMS) vectors are generated. This study uses the value of 5 as n after several trials. In addition, the number of identical vectors in the HM is limited to prevent the premature HM which has many local optima. This study uses the value of 2 as the number of maximum allowed identical vectors because it was frequently used in the literature.

For pursuing the diversity of vectors in the HM, if a parcel contains a species which is not yet included in the HM, it is deterministically selected regardless of Equation 10 every tenth iteration.

For the MCSP, the HS algorithm approached 24 different P cases (P = 1, 2, …, 24) as shown in Table 1. The HS results were also compared with those of another meta-heuristic algorithm, SA [2] as well as those of exact method [3]. When compared with SA, HS found better solutions in 14 cases (P = 8, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24) while it found worse solution only once (P = 9). When compared with exact method, HS found optimal solutions or near-optimal solutions with maximum one species gap.

In each P case, the HS used various HMS values {10, 30, 50, 100} and P_{Memory} values {0.9, 0.95, 0.97}. Here, P_{Random} is $(1-P_{Memory})$.

Table 2 shows the results with various algorithm parameter values when P = 24. HS found the best solution (426) four times out of 12 cases, ranging 421 to 426. The last column shows the iteration at which the HS found the solution. This study tried up to 30,000 iterations, taking up to 34 minutes per each run with MS Excel.

Table 1. Comparison of MCSP Results

P	HS	SA	Opt
1	254	254	254
2	318	318	318
3	356	356	356
4	374	374	374
5	384	384	384
6	390	390	390
7	395	395	395
8	400	398	400
9	402	403	403
10	405	405	406
11	408	407	408
12	410	409	410
13	412	411	412
14	414	413	414
15	416	415	416
16	417	417	418
17	419	418	419
18	420	419	420
19	421	420	422
20	422	421	423
21	423	422	424
22	424	423	425
23	425	424	426
24	426	425	NA
25	NA	426	NA

Table 2. Results with Various Parameter Values

HMS	P_{Memory}	Solution	Iteration
10	0.90	425	6,783
	0.95	424	17,298
	0.97	421	1,557
30	0.90	425	24,220
	0.95	426	5,675
	0.97	424	9,049
50	0.90	425	15,336
	0.95	425	4,356
	0.97	426	10,266
100	0.90	425	28,943
	0.95	426	26,389
	0.97	426	10,778

Table 3 shows alternative solutions the HS found (P = 24 and # of covered species = 426). If certain reserve is not available in one solution, other alternative solutions that do not contain the specific reserve can be considered. While interchange heuristic [3] found three alternative solutions, the HS found 25 alternative solutions.

Table 3. Alternative Solutions of HS

Solutions (P = 24, # of covered species = 426)
9 11 24 26 55 75 102 114 120 136 147 169 222 224 289 314 319 324 345 357 367 375 428 440
9 11 24 26 55 75 102 120 128 136 147 169 222 224 289 314 319 324 345 357 367 375 428 440
9 11 24 26 55 75 102 120 136 147 169 222 224 252 289 314 319 324 345 357 367 375 428 440
9 11 24 26 55 75 102 120 136 147 169 222 224 268 289 314 319 324 345 357 367 375 428 440
9 11 24 27 55 75 120 121 134 147 169 222 224 252 289 313 314 319 324 364 375 376 428 440
9 11 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 364 367 375 376 428 440
9 11 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 364 375 376 386 428 440
9 11 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 364 375 376 403 428 440
9 11 24 27 55 75 120 121 135 147 169 222 224 252 289 313 314 319 324 364 375 376 428 440
9 11 24 27 55 75 120 121 135 147 169 222 224 252 289 314 319 324 364 375 376 386 428 440
9 24 26 55 75 120 121 136 147 148 169 222 224 278 289 314 319 324 357 364 367 375 428 440
9 24 26 55 75 120 121 136 147 148 169 222 224 278 289 314 319 324 364 367 375 377 428 440
9 24 27 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 345 357 375 386 428 440
9 24 27 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 357 364 375 386 428 440
9 24 27 55 75 120 121 134 147 169 222 224 252 289 313 314 319 324 347 364 375 376 428 440
9 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 347 364 367 375 376 428 440
9 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 347 364 375 376 386 428 440
9 24 27 55 75 120 121 134 147 169 222 224 252 289 314 319 324 347 364 375 376 403 428 440
9 24 27 55 75 120 121 135 147 169 222 224 252 289 313 314 319 324 347 364 375 376 428 440
9 24 27 55 75 120 121 135 147 169 222 224 252 289 314 319 324 347 364 367 375 376 428 440
9 24 44 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 345 357 375 386 428 440
9 24 45 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 345 357 375 386 428 440
9 24 46 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 345 357 375 386 428 440
9 24 46 55 75 114 120 121 135 141 147 169 224 278 289 314 319 324 357 364 375 386 428 440
9 24 55 63 75 114 120 121 135 141 147 169 224 278 289 314 319 324 345 357 375 386 428 440

ISSN: 1790-5117

ISBN: 978-960-6766-47-3

5 Conclusions

The HS algorithm was applied to an ecological optimization problem (MCSP) where the number of preserved species in an area is to be maximized while limiting the number of considered area parcels.

The HS model in this study was modified from its original structure in order to be applied to MCSP:

- Because the candidate value of each variable is 0 or 1, the HS model does not have pitch adjustment operation.
- Because each solution vector is sparse (many 0's and few 1's), the chance to be selected is limited (much lower than 50%) based on the number of species in each area parcel.
- In order to start with good environment, solution vectors in initial HM are chosen from n-times generations.
- For the diversity of solution vectors in HM, a vector which has new species is deterministically included in the HM.

The HS model was applied to real-world problem (the state of Oregon) which has 426 species and 441 parcels. In 24 cases with different parcel number, the HS found global optimum solutions in 15 cases and near-optimal solutions (only one species gap) 9 cases.

When compared with other meta-heuristic algorithm, the HS algorithm found better solutions than those of the SA algorithm in 14 cases while the former found worse solution only once.

Another advantage of the HS algorithm is the fact that it suggests many alternative solutions because it simultaneously handles multiple solution vectors. For example, the HS found 25 different alternative solutions for the case of $P = 24$.

If HS considers additional problem-specific heuristics as well as basic operators, the authors expect that it will perform better than the proposed HS model in this study.

References:

[1] ReVelle, C., Williams, J. C., and Boland, J. J., "Counterpart models in facility location science and reserve selection science," Environmental Modelling and Assessment, 7, 71-80, 2002.

[2] Csuti, B., S. Polasky, P. H. Williams, R. L. Pressey, J. D. Camm, M. Kershaw, A. R. Kiester, B. Downs, R. Hamilton, M. Huso, and K. Sahr. "A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon," Biological Conservation, 80, 83-97, 1997.

[3] Rosing, K. E., ReVelle, C. S., and Williams, J. C., "Maximizing Species Representation under Limited Resources: A New and Efficient Heuristic," Environmental Modeling and Assessment, 7(2), 91-98, 2002.

[4] Mizumori, M. M., ReVelle, C. S., and Williams, J. C., "The maximal multiple-representation species problem solved using heuristic concentration", in The Next Wave in Computing, Optimization, and Decision Technologies, B. L. Golden, S. Raghavan, and E. A. Wasil, eds., Springer: New York, 183-198, 2005.

[5] Geem, Z. W., Kim, J. H., and Loganathan, G. V., "A New Heuristic Optimization Algorithm: Harmony Search," Simulation, 76(2), 60-68, 2001.

[6] Geem, Z. W., Tseng, C. -L., and Park, Y., "Harmony Search for Generalized Orienteering Problem: Best Touring in China," Lecture Notes in Computer Science, 3612, 741-750, 2005.

[7] Geem, Z. W. and Choi, J. Y., "Music Composition Using Harmony Search Algorithm," Lecture Notes in Computer Science, 4448, 593-600, 2007.

[8] Geem, Z. W., "Harmony Search Algorithm for Solving Sudoku," Lecture Notes in Artificial Intelligence, 4692, 371-378, 2007.

[9] Geem, Z. W., "Optimal Cost Design of Water Distribution Networks using Harmony Search," Engineering Optimization, 38(3), 259-280, 2006.

[10] Geem, Z. W., "Optimal Scheduling of Multiple Dam System Using Harmony Search Algorithm," Lecture Notes in Computer Science, 4507, 316-323, 2007.

[11] Geem, Z. W., Lee, K. S., and Park, Y., "Application of Harmony Search to Vehicle Routing," American Journal of Applied Sciences, 2(12), 1552-1557, 2005.

[12] Lee, K. S., Geem, Z. W., Lee, S. -H., Bae, K. -W., "The Harmony Search Heuristic Algorithm for Discrete Structural Optimization," Engineering Optimization, 37(7), 663-684, 2005.

[13] Geem, Z. W., "Novel Derivative of Harmony Search Algorithm for Discrete Design Variables," Applied Mathematics and Computation, In Press.