Comparison of HSRNAFold and RNAFold Algorithms for RNA Secondary Structure Prediction

Abdulqader M. Mohsen
School of Computer Sciences
Universiti Sains Malaysia (USM)
Penang, Malaysia 11800
Email: amayemen@yahoo.com

Ahamad Tajudin Khader
School of Computer Sciences
Universiti Sains Malaysia (USM)
Penang, Malaysia 11800
Email: tajudin@cs.usm.my

Dhanesh Ramachandram
School of Computer Sciences
Universiti Sains Malaysia (USM)
Penang, Malaysia 11800
Email: dhaneshr@cs.usm.my

***Abstract*—Ribonucleic Acid (RNA) has important structural and functional roles in the cell and plays roles in many stages of protein synthesis. The structure of RNA largely determines its function. Current physical methods for structure determination are time-consuming and expensive, thus the methods for the computational prediction of structure are necessary. Various algorithms that have been used for RNA structure prediction based in minimum free energy include dynamic programming (DP) and meta heuristic algorithms. One of the most recent meta heuristic algorithms is Musician's behavior-inspired harmony search (HS) algorithm that has been successful in numerous complex optimization problems. This paper builds on the previous work of the harmony search algorithm (HSRNAFold) which was used to find the RNA secondary structure with minimum free energy. In this paper, the accuracy of prediction is compared to the dynamic programming technique RNAFold. The results show that HSRNAFold is able to predict more accurate structures than RNAFold for all test sequences.**

***Index Terms*—Meta heuristic algorithms; Dynamic programming algorithms;RNA folding; Minimum Free Energy.**

I. Introduction

RNA is nucleic acid consist of a long linear polymer of nucleotide units found in the nucleus. RNA is similar to DNA but usually consists of a single strand instead of double-stranded, containing ribose rather than deoxyribose, and has the base uracil (U) in place of thymine (T). There are a various forms of RNA are found: heterogeneous nuclear RNA (hnRNA), messenger RNA (mRNA), transfer RNA tRNA), ribosomal RNA (rRNA), and small nuclear RNA. Structurally, hnRNA and mRNA are both single stranded, while rRNA and tRNA form three-dimensional molecular configurations. Each type of RNA has a different role in various cellular processes such as carrying genetic information (mRNA), interpreting the code (rRNA), and transferring genetic code (tRNA). It also performs different functions which include: catalyzing chemical reactions [1], [2], directing the site specific modification of RNA nucleotides, controlling gene expression, modulating protein expression and serving in protein localization [3].

The function of RNA molecules determines many diseases caused by RNA viruses. Identifying the secondary structure of an RNA molecule is the fundamental key to understand its biological function [4],[5].

The physical methods to determine the secondary structure such as x-ray diffraction and NMR spectroscopy are difficult, error prone, time-consuming and expensive. Therefore, computational approaches are appropriate alternatives to predict the secondary structure of an RNA molecule. Two different computational approaches exist for RNA secondary prediction; either comparative methods or thermodynamic optimization.

Since RNA folding is subject to the laws of thermodynamics, there is an assumption that the correct structure is a low energy structure [6]. The stability of the secondary structure depends on the amount of free energy released to form the base pairs. Thus, the more negative the free energy of a structure is, the more stable a particular sequence is formed. This structure is called the minimum free energy (MFE) secondary structure [7]. In case of only the sequence of a given RNA molecule is known, the ab initio methods are used to perform RNA secondary structure prediction as an energy minimization problem through the use of thermodynamic models. These methods are either dynamic programming or meta-heuristics.

Mfold [8] and *RNAFold* [9]algorithms are DP techniques for RNA secondary structure prediction based on finding the minimum free energy. Dynamic programming algorithm as mathematical technique can hit the global optima in small problems. But in real world problems there are some drawbacks. For examples, when the number of variables increases, the number of evaluations of the recursive function will also increase exponentially. For RNA Secondary structure prediction the large number of structure alternatives make it difficult to determine which one is more correct [10]. In addition, these algorithms only predict the minimum free energy structure, while the native structure usually has a free energy of about 5-10% from the minimum free energy of the sequence.

On the other hand, many meta-heuristics algorithms were proposed such as Genetic Algorithm; *RnaPredict* [11],[12], Simulated Annealing; *SARNA-Predict* [10], Particle Swarm Optimization; *HelixPSO* [13] and Harmony search algorithm; *HSRNAFold* [14]. These algorithms generate a structure vector instead of only the minimum free energy structure.
This paper builds on the previous research in harmony search [14] which showed that it is possible to use an HS to minimize the free energy associated with RNA secondary structures and predict structures similar to the known structures. The performance of *HSRNAFold* is compared to dynamic programming algorithm *RNAFold* based on standard sets of RNA test

molecules.

II. RNA SECONDARY STRUCTURE PREDICTION

RNA molecule consists of a single stranded sequence of four nucleotides: adenine (A), guanine (G), cytosine (C), or uracil (U). This linear sequence is the primary structure of RNA molecule.

The RNA strand can fold back upon itself. During the folding process, the hydrogen bonds between the different nucleotides form the base pairs. These hydrogen bonds which occur mostly between G and C, or A and U are called the Watson-Crick base pairs and the bond between G and U is called the wobble base pair. These base pairs; GC, AU, and GU, and their mirrors, CG, UA, and UG are called the canonical base pairs.

RNA secondary structure is defined by a set of base pairs that satisfy the following constraints: i) for (i,j), it must be canonical base pairs; ii) each base cannot share more than one base; iii) pairing bases must be at least three bases apart $i-j > 3$; and iv) Two base pairs must not cross, i.e.:$i, j \bigcap i^{'}, j^{'} = \Phi$ or for all (i, j), $(i^{'}, j^{'})$ either $i < i^{'} < j^{'} < j$ or $i^{'} < i < j < j^{'}$ holds.

The stability of the RNA secondary structure is quantified as the amount of free energy being released or used by the forming base pairs. The stability increases according to the number of GC versus AU and GU base pairs and the number of base pairs in a hairpin loop region. The number of unpaired bases decreases the stability of the structure such as interior loops, hairpin loop or bulges.

III. APPROACH

A. harmony search algorithm

HS algorithm is a meta-heuristic searching algorithm inspired by the music improvisation process in which the musicians in an orchestra try to find a fantastic harmony through musical improvisations. In the improvisation processes, musicians look for the best combination of state in order to produce a fantastic harmony just like optimization process seeks a best solution (global optimum) determined by objective function evaluation.

HS algorithm is a stochastic iterative search optimization algorithm based on successive update steps. Geem et al. [15] in 2001 were the first to apply harmony search to optimization problems.

This algorithm was applied to many minimization problems such as: continuous engineering optimization, vehicle routing, combined heat and power economic dispatch, water pump switching problem, optimal scheduling of multiple dam system and transport energy modeling [16].

In general, HS has five steps: i) Initializing the problem and algorithm parameters including a representation of solution vectors to the problem; ii) creating an initial harmony memory (HM) of candidate solutions as a solution vector ; iii) improvising a new harmony from HM; iv) updating the harmony memory; v) and ,finally, checking the stopping criterion.

HS has four main parameters that direct the search toward the most favorable areas of the search space. These parameters are:

- Harmony memory size (HMS) represents the total number of harmonies in the HM.
- Harmony memory consideration rate (HMCR) represents the probability of picking up values from HM to the variables in the solution vector.
- Random selection rate (RSR) represents the probability of randomly choosing feasible values from the rage of all possible values to the variables in the solution vector.
- Pitch adjusting rate (PAR) represents the probability of further adjusting the pitch with neighboring pitches.

HS manages harmony memory vectors of harmonies in which each harmony represents a potential solution to the given problem. Harmony memory size (HMS) represents the total number of harmonies in the HM. Each harmony in HM is evaluated to determine its relative fitness within the harmony memory vectors. In each cycle or 'improvisation',a new harmony retained via each improvisation process. HS parameters: harmony memory consideration rate ,random selection rate , and pitch adjusting rate , are applied to the HM for each improvisation process to guide the algorithm to the promising area in the search space.

After the improvisation of a new harmony is complete, the new harmony is evaluated by its objective function (fitness function). If the value of its objective function is better than the value of the objective function of the worst harmony in the HM, the new harmony is included in the HM and the existing worst harmony is excluded from the HM. Consequently, the vectors are sorted out based on their fitness functions.
Finally, the cycle repeats itself with a new harmony. After a varying number of improvisations, the algorithm will converge

Algorithm 1 HSRNAFold algorithm

```
Generate all possible base-pairs pool.
Generate all possible helices from the base-pairs pool.
Initial HS parameters: HMCR, RSR and PAR
for i = 1 to HMS do
    Generate feasible structure[i] randomly from the helices
    FreeEnergy[i] = Evaluate(structure[i])
    Sort the generated structures according to the freeEnergy
end for
for i = 1 to number_of_iterations do
    Change the rate of HMCR
    Change the Rate of PAR
    Generate new feasible structure depending on the HMCR,
    RSR and PAR
    newFeeEnergy = Evaluate(newstructure);
    if newFeeEnergy < FreeEnergy[HMS] then
        structure[HMS] = newstructure
    end if
end for
Return the best structure
```

TABLE I
TESTING RNA ORGANISMS WITH THEIR ACCESSION NUMBERS, CLASSES, LENGTHS AND NUMBER OF BASE PAIRS IN THEIR KNOWN STRUCTURES.

Organism	Accession Number	RNA Class	Length	BP in Known Structure
Geobacillus stearothermophilus	AJ251080	5S rRNA	117	38
Saccharomyces cerevisiae	X67579	5S rRNA	118	37
Escherichia coli	V00336	5S rRNA	120	40
Haloarcula marismortui	AF034620	5S rRNA	122	38
Thermus Aquaticus	X01590	5S rRNA	123	40
Deinococcus radiodurans	AE002087	5S rRNA	124	40
Metarhizium anisopliae var. anisopliae(3)	AF197120	Group I intron, 23S rRNA	394	120
Chlorella saccharophila	AB058310	Group I intron,16S rRNA	454	126
Hildenbrandia rubra	L19345	Group I intron, 16S rRNA	543	141
Acanthamoeba grifini	U02540	Group I intron, 16S rRNA	556	131
Drosophila virilis	X05914	16S rRNA	784	233
Xenopus laevis	M27605	16S rRNA	945	254

to the best harmony which represents a sub optimal or optimal solution to the given problem.

B. HSRNAFold

HSRNAFold is a meta-heuristic algorithm based on harmony search for RNA secondary structure prediction by minimizing the the free energy of a particular RNA sequence [14].

HSRNAFold starts by creating a pool of all possible pairs, and consequently stems (helices) that can be generated from the given sequence using helix generation algorithm described in [12],[10],[13]. This algorithm will fold the sequence into a structure and also reduce the search space.

After building the set of helices, H, the algorithm tries to find a subset of H that defines an optimal secondary structure with the minimum free energy [12],[10],[13]. Thus, the structure prediction becomes a combinatorial optimization problem of picking a subset x from H. Since RNA folds into a structure with near minimal free energy, *HSRNAFold* attempts to find the combination of helices that produce a feasible structure with the lowest possible free energy.

In *HSRNAFold*, Each harmony in the HS encodes a potential RNA feasible structure. A permutation-based representation is used to encode the RNA secondary structure. Each helix in H is numbered by an integer ranging from 0 to n ; where n is the number of helices in H. For example, if $n = 4$, $\{2, 3, 1, 0\}$ and $\{1, 0, 3, 2\}$ will be two possible permutations. To produce feasible structures, each permutation is decoded from left to right. Each helix is checked for conflicts with helices to its left. If there are no conflicts found, the helix is retained; otherwise it is discarded. After that, the algorithm generates a new structure based on the memory consideration rate, pitch adjustment and random selection.

Then, the new structure is evaluated by calculating its free energy; if it is better than the worst structure in the HM, then it will be included in the HM and the existing worst structure will be excluded.

Finally, the algorithm iterates through the solution vectors and generates a new structure in each iteration until the structure of the minimum free energy is found or the maximum number of iterations is reached. Algorithm 1 shows the pseudo code of the proposed *HSRNAFold* algorithm

IV. RESULTS AND DISCUSSION

All *HSRNAFold* experiments were conducted on an Intel Core 2 Quad. Each CPU is a 2.4 GHz. *HSRNAFold* was implemented using C# 2005. Twelve sequences were used for testing *HSRNAFold*.

All sequences were taken from the Comparative RNA web site [17], and their details are summarized in Table I.

It should be mentioned that the names of the organisms are used to refer to these specific RNA sequences in this manuscript.

The parameters which remained fixed throughout all experiments presented here are as follows: HMS*100 iterations, a HMCR probability (P_{HMCR}) varies between 0.70 and 0.90 and a PAR probability (P_{PAR}) varies between 0.1 and 0.5.

The *RNAFold* Web server was used to generate the *RNAFold* results presented here with default settings.

Six measures are used for a comparison of *HSRNAFold* to *RNAFold* which can be defined as follows: true positive (TP) represents the number of base pairs for both the known and predicted structure; false positive (FP) is the total of predicted base pairs which are not found in the known structure; false negative (FN) indicates the base pairs in the known structures which have not been predicted. Sensitivity (SE) is the ratio between TP and the total number of base pairs found in the native structure. Specificity (SP) is the ratio between TP and all base pairs predicted. Finally, F-measure (FM) is a metric that combines both sensitivity and specificity into a single performance measure: $FM = 2 * SE * SP/(SE + SP)$.

In table II, the results for *HSRNAFold* (HS) are compared to those of the minimum free energy structure as calculated by the *RNAFold* (Fold) algorithm of the ViennaRNA package. For TP, it is clearly that *HSRNAFold* is able to predict TP better than *RNAFold* in seven out of twelve sequences. They were equal in only *Saccharomyces cerevisiae* sequence. On average, *HSRNAFold* predicted 52.75 percent TP base pairs which is better than *RNAFold* (45.17percent). *HSRNAFold* predicted a fewer

TABLE II
COMPARISON OF BEST RESULTS OF HSRNAFOLD AND THE MINIMUM FREE ENERGY STRUCTURES OF RNAFOLD (FOLD) IN TERMS OF CORRECTLY PREDICTED BASE PAIRS (TP), INCORRECTLY PREDICTED BASE PAIRS (FP) AND BASE PAIRS IN NATIVE STRUCTURE THAT WERE NOT PREDICTED (FN); THE BEST RESULTS ARE IN BOLD.

Organism	Seq. Length	HS	Fold	HS	Fold	HS	Fold
		TP	TP	FP	FP	FN	FN
Geobacillus stearothermophilus	117	**23**	12	**7**	23	**15**	26
Saccharomyces cerevisiae	118	**33**	**33**	**4**	8	**4**	4
Escherichia coli	120	**28**	10	**6**	28	**12**	30
Haloarcula marismortui	122	**29**	16	**5**	23	**9**	22
Thermus Aquaticus	123	26	**27**	**3**	11	14	**13**
Deinococcus radiodurans	124	26	**31**	**4**	5	14	**9**
Metarhizium anisopliae var. anisopliae(3)	394	69	**70**	**30**	55	51	**50**
Chlorella saccharophila	454	84	**103**	50	**49**	42	**23**
Hildenbrandia rubra	543	**84**	53	**63**	123	**61**	88
Acanthamoeba grifini	566	**90**	76	**40**	102	**41**	55
Drosophila virilis	784	**56**	37	**147**	221	**177**	196
Xenopus laevis	945	**85**	74	**109**	173	**169**	180
Average		**52.75**	45.17	**39**	68.42	**50.75**	58

TABLE III
COMPARISON OF BEST RESULTS OF HSRNAFOLD AND THE MINIMUM FREE ENERGY STRUCTURES OF RNAFOLD (FOLD) IN TERMS OF SENSITIVITY (SE), SPECIFICITY (SP), AND F- MEASURE (FM); THE BEST RESULTS ARE IN BOLD.

Organism	Seq.	HS	Fold	HS	Fold	HS	Fold
	Length						
		SE	SE	Sp	Sp	FM	FM
Geobacillus stearothermophilus	117	**60.53**	31.58	**76.67**	34.29	**67.65**	32.88
Saccharomyces cerevisiae	118	**89.19**	89.19	**89.19**	80.49	**89.19**	84.62
Escherichia coli	120	**70**	25	**82.35**	26.32	**75.68**	25.64
Haloarcula marismortui	122	**76.32**	42.11	**85.29**	41.03	**80.56**	41.56
Thermus Aquaticus	123	65	**67.5**	**89.66**	71.05	**75.36**	69.23
Deinococcus radiodurans	124	65	**77.5**	**86.67**	86.11	74.29	**81.58**
Metarhizium anisopliae var. anisopliae(3)	394	57.5	**58.33**	**69.7**	56	**63.01**	57.14
Chlorella saccharophila	454	66.67	**81.75**	62.69	**67.76**	64.62	**74.1**
Hildenbrandia rubra	543	**59.57**	37.59	**57.93**	30.11	**58.74**	33.44
Acanthamoeba grifini	566	**68.7**	58.02	**69.23**	42.7	**68.97**	49.19
Drosophila virilis	784	**24.03**	15.88	**27.59**	14.34	**25.69**	15.07
Xenopus laevis	945	**33.46**	29.13	**43.18**	29.96	**37.95**	29.54
Average		**61.33**	51.13	**70.01**	48.35	**65.14**	49.5

false positive base pairs than *RNAFold* in all sequences except *Chlorella saccharophila* sequence. Furthermore, the average performance of *HSRNAFold* clearly outperformed *RNAFold* in terms of FP (39 percent and 68.42 percent respectively). For FN, *HSRNAFold* failed to predict a fewer number of base pairs, as in the known structures, than *RNAFold* in seven sequences. On average performance, *HSRNAFold* algorithm substantially outperformed *RNAFold* by 7.25 percent.

Compared to *RNAFold* in terms of sensitivity, specificity and F-measure, table III reveals that *HSRNAFold* achieved substantially a higher sensitivity than *RNAFold* in eight out of twelve sequences. On the overall average, it is clear that *HSRNAFold* outperformed *RNAFold* in terms of sensitivity by 10.20 percent. *HSRNAFold* has a higher specificity in eleven out of twelve sequences except *Chlorella saccharophila* sequence. The overall performance shows that *HSRNAFold* was significantly better than *RNAFold* by 21.67 percent. Finally, for the FM, *HSRNAFold* produced better results than *RNAFold* in ten out of twelve sequences except *Metarhizium anisopliae var. anisopliae(3)* and *Chlorella saccharophila* sequences. The average performance of HSRNAFold is better than that of *RNAFold* by 15.64 percent.

It has been noted that the *HSRNAFold* is clearly outperformed *RNAFold* in all sequences larger than 500 nucleotides with all measures.

Fig.1 shows secondary structure for three RNA sequences *Geobacillus stearothermophilus*, *Saccharomyces cerevisiae* and *Escherichia coli*. jViz.Rna [18] was used to create the structures' images. The first column shows the secondary structures predicted by *RNAFold*, the middle column shows the secondary structures predicted by *HSRNAFold* and last column shows the native structure for these three sequences as in the comparative RNA web site [17]. Accordingly, the first row represents the structures of *Geobacillus stearothermophilus*, the second row represents the structures of *Saccharomyces cerevisiae* whereas the last row represents *Escherichia coli* structures.

For *Geobacillus stearothermophilus*, the structure predicted by *HSRNAFold* (Fig. 1b) is somewhat similar to the native fold (Fig. 1c). On contrast, there is no real similarity between

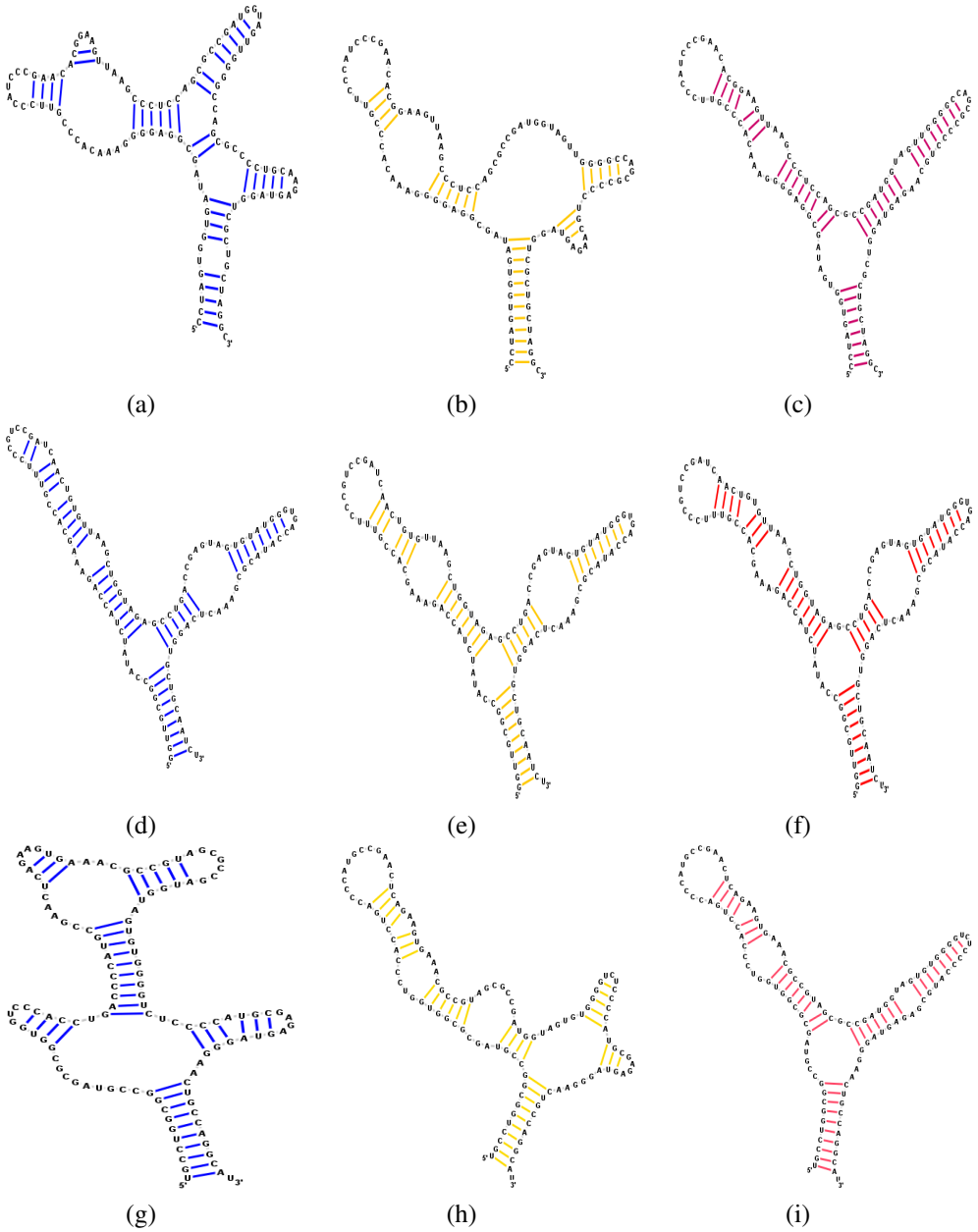

Fig. 1. HSRNAFold best structures of G. stearothermophilus 5S rRNA (AJ251080),S. cerevisiae 5S rRNA (X67579) and E. coli 5S rRNA(V00336) compared with the MFE structures predicted by RNAFold with regard to native structures. (a) G. stearothermophilus (RNAFold). (b) G. stearothermophilus (HSRNAFold). (c) G. stearothermophilus (Native). (d) S. cerevisiae (RNAFold). (e) S. cerevisiae (HSRNAFold). (f) S. cerevisiae (Native). (g) E. coli(RNAFold). (h) E. coli (HSRNAFold). (i) E. coli (Native).

the structure predicted by *RNAFold* (Fig. 2a) and the native one (Fig. 2c). Similarly, for *Escherichia coli*, the structure predicted by *HSRNAFold* (Fig. 1h) is similar to the native fold (Fig. 1i) with some exceptions where as the structure predicted by *RNAFold* (Fig. 1g) is different to a large extent. These exceptions can be observed, especially, on the right most branch and in the multi branch loop.

For *Saccharomyces cerevisiae*, the both structures predicted by *RNAFold* (Fig. 1d) and *HSRNAFold* (Fig. 1e) are similar to each other and to the native fold (Fig. 1f) with small differences can still be observed. Particularly, *HSRNAFold* has the superiority of similarity to the native fold. For example, there is a multi branch loop and three internal loops in both structures. The stems look very similar in all three branches with the exceptions. These exceptions resulted because the inability of the current thermodynamic models to predict the two CU non-canonical base pairs.

V. Conclusion

This paper builds on the previous research in harmony search; *HSRNAFold*. Comparisons performed to *RNAFold* de-

termined that *HSRNAFold* can predict structures with a higher number of true positive base pairs and a lower number of false positive base pairs than the minimum $\triangle G$ structure computed by *RNAFold* for the majority of sequences tested.

Overall, the prediction accuracy of both methods is good for shorter sequences. Prediction accuracy decreases as sequence lengths increase; this is a result of limitations in the performance of the thermodynamic models.

In future work, Code optimization, effect of parameters adaptation and the possibility to hybridize HS with other optimization algorithms to enhance the HS performance should be studied and modeling non-canonical base pairs to further improve the prediction accuracy of *HSRNAFold*.

ACKNOWLEDGMENT

The first author would like to acknowledge University of Science Malaysia for the financial support under USM Fellowship scheme. Also he would like to acknowledge the Ministry of Science, Technology and Innovation, MOSTI for funding this research under the EscienceFund grant, for "An Intelligent Genetic Algorithms' Parameters Adaptation for RNA Folding Problem" Project [USM0001306].

REFERENCES

[1] J. A. Doudna and T. R. Cech, "The chemical repertoire of natural ribozymes," *Nature*, vol. 418, no. 6894, pp. 222–228, 2002.

[2] J. L. Hansen, T. M. Schmeing, P. B. Moore, and T. A. Steitz, "Structural insights into peptide bond formation," in *Proceedings of the National Academy of Sciences*, vol. 99, 2002, pp. 11 670–11 675.

[3] J. Bachellerie, J. Cavaill, and A. Httenhofer, "The expanding snorna world," *Biochimie*, vol. 84, no. 8, pp. 775–790(16), August 2002.

[4] H. Tsang and K. Wiese, "Sarna-predict: A study of rna secondary structure prediction using different annealing schedules," in *Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, 2007, pp. 239–246.

[5] M. Neethling and A. Engelbrecht, "Determining rna secondary structure using set-based particle swarm optimization," in *IEEE Congress on Evolutionary Computation(CEC2006)*, 2006, pp. 1670–1677.

[6] A. A. Deschenes, "A genetic algorithm for rna secondary structure prediction using stacking energy thermodyamic models," Ph.D. dissertation, Simon Fraser University, 2005.

[7] D. M. Layton and R. Bundschuh, "A statistical analysis of rna folding algorithms through thermodynamic parameter perturbation," *Nucleic Acids Research*, vol. 33, no. 2, p. 519524, 2005.

[8] M. Zuker, "Mfold web server for nucleic acid folding and hybridization prediction," *Nucleic Acids Research*, vol. 31, no. 13, p. 34063415, 2003.

[9] I. Hofacker, W. Fontana, P. Stadler, S. Bonhoeffer, M. Tacker, and P. Schuster, "Fast folding and comparison of rna secondary structures (the vienna rna package)," *Monatshefte f. Chemie*, vol. 125, pp. 167–188, 1994.

[10] H. H. Tsang, "Sarna-predict: A permutation-based simulated annealing algorithm for rna secondary structure prediction," Ph.D. dissertation, Simon Fraser University, 2007.

[11] K. Wiese, A. Deschnes, and A. Hendriks, "Rnapredict - an evolutionary algorithm for rna secondary structure prediction," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 5, no. 1, pp. 25–41, 2007.

[12] A. Hendriks, "A parallel evolutionary algorithm for rna secondary structure prediction," Ph.D. dissertation, Simon Fraser University, May2005.

[13] M. Geis and M. Middendorf, "A particle swarm optimizer for finding minimum free energy rna secondary structures," in *Swarm Intelligence Symposium, 2007. SIS 2007. IEEE*, 2007, pp. 1–8.

[14] A. M. Mohsen, A. T. Khader, and D. Ramachandram, "Hsrnafold: A harmony search algorithm for rna secondary structure prediction based on minimum free energy," in *Innovations in Information Technology, 2008. IIT 2008.*, 2008, pp. 11–15.

[15] Z. W. Geem, J. H. Kim, and G. Loganathan, "A new heuristic optimization algorithm: Harmony search," *SIMULATION*, vol. 76, no. 2, pp. 60–68, 2001.

[16] M. Mahdavi, M. Fesanghary, and E. Damangir, "An improved harmony search algorithm for solving optimization problems," *Applied Mathematics and Computation*, vol. 188, no. 2, pp. 1567–1579, 2007.

[17] J. J. Cannone, S. Subramanian, M. N. Schnare, J. R. Collett, L. D'Souza, L. N. Du Y, Feng B, M. L.V, M. K.M, P. N, Y. N. Shang Z, and G. R.R., "The comparative rna web (crw) site: an online database of comparative sequence and structure information for ribosomal,intron, and other rnas," *BMC Bioinformatics*, vol. 3, no. 35, pp. 169–172, 2002.

[18] K. C. Wiese, E. Glen, and A. Vasudevan, "jviz.rna - a java tool for rna secondary structure visualization," *IEEE Transactions on NanoBioscience*, vol. 4, no. 3, pp. 212–218, 2005.