

Alignment of Noncoding Ribonucleic Acids with Pseudoknots Using Context-Sensitive Hidden Markov Model

Abstract

Up to now, various signal processing techniques have been used to predict protein-coding genes that are unsuitable for predicting ribonucleic acids (RNAs). Modeling a gene network can be employed in various fields, such as the discovery of new drugs, reducing the side effects of treatment methods, further identifying genetic diseases and treatments for genetic disorders by influencing the activity of effectual genes, preventing the growth of unwanted tissues via growth weakening and cell reproduction, and also for many other applications in the fields of medicine and agriculture. The main purpose of this study was to design a suitable algorithm based on context-sensitive hidden Markov models (csHMMs) for the alignment of secondary structures of RNAs, which can identify noncoding RNAs. In this model, several RNA families are compared, and their existing similarities are measured. An expectation–maximization algorithm is used to estimate the model's parameters. This algorithm is the standard algorithm to maximize HMM parameters. The alignment results for RNAs belonging to the hepatitis delta virus family showed an accuracy of 83.33%, a specificity of 89%, and a sensitivity of 97%, and RNAs belonging to the purine family showed an accuracy of 65%, a specificity of 76%, and a sensitivity of 76%. The results show that csHMMs, in addition to aligning the primary sequences of RNAs, would align the secondary structures of RNAs with high accuracy.

Keywords: Context-sensitive hidden Markov models, expectation–maximization algorithm, noncoding ribonucleic acids, structural alignment

Submitted: 09-Feb-19

Revision: 22-Mar-19

Accepted: 15-May-19

Published: 24-Oct-19

Introduction

Different cellular mechanisms that ensure living organisms are the result of effective cooperation biomolecules such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and proteins. For a long time, it was thought that these molecules are the proteins that further structural change and chemical reactions and are responsible for cells' regulatory functions. Accordingly, the DNA to the protein coding information storage and RNA molecules as an intermediary between DNA and protein were seen. However, some recent observations in the field of molecular biology show that the traditional approach to explaining many biological actions in complex multicellular organisms, such as plants, insects, and animals, is imperfect and limited.

New studies on the genomes of RNA in the presence of numerous noncoding DNAs revealed that they are not translated into proteins but act directly as RNA plays an important role in various biological actions.^[1] In addition to the examples already known, such as transfer RNA and ribosomal RNA, the number of known operational noncoding RNA (ncRNAs) is abundant, and their performance is extremely diverse. Because RNAs based on a sequence are capable of direct confrontations with other RNA and DNA molecules, they can be useful for regulatory mechanisms including identifying specific nucleotide sequences.^[2]

ncRNA are small noncoding sequences involved in the gene expression regulation of many biological processes and diseases.^[3] Recent developments show that ncRNAs play an important role in gene silencing, RNA processing and modification, control transcription and translation, and sustain many other regulatory functions.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Bakhshayesh NM, Shamsi M, Sedaaghi MH, Ebrahimnezhad H. Alignment of noncoding ribonucleic acids with pseudoknots using context-sensitive hidden Markov model. *J Med Sign Sens* 2019;9:252-8.

**Nayer Mostaghim Bakhshayesh¹,
Mousa Shamsi¹,
Mohammad Hossein Sedaaghi²,
Hossein Ebrahimnezhad²**

¹Faculty of Biomedical Engineering, Sahand University of Technology, Tabriz, Iran, ²Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran

Address for correspondence:
Dr. Mousa Shamsi,
Faculty of Biomedical Engineering, Sahand University of Technology, Tabriz, Iran.
E-mail: shamsi@sut.ac.ir

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.JMSS_11_19

Quick Response Code:



The structure of RNA molecule affects its function; thus, many ncRNAs have preserved a secondary structure that uses the position of the ncRNA genes. The structure has been preserved in such a way that some functional RNAs with the same applications' functionality have similar secondary structures, so the structure in the RNA is well preserved and remains stable. Therefore, search methods should enable modeling these structures.

One of the operations that is performed on RNA is a structural comparison. In recent years, thousands of sequencing projects around the world have been creating enormous volumes of RNA data, which has led to the discovery and description of a rapidly increasing number of ncRNAs in eukaryotic genomes.^[4] Comparing the structure of two or more RNAs has many applications in structure prediction and in finding a motif. To model the RNAs while preserving their secondary structures, statistical models will be needed to describe the correlation between base pairs. So far, numerous statistical models to provide targeted RNA secondary structures only use a limited number of RNA secondary structure classes.

Alignment is another operation that compares DNA, RNA, and amino acid sequences. Generally, the alignment of two sequences finds the most similarity between the two sequences so that one of the sequences is probably achieved from the deletion, insertion, or substitution of the nucleotide in another sequence.^[5] Many search methods such as Basic Local Alignment Search Tool (BLAST), Fast Alignment (FASTA), and Database of Protein Domains, Families and Functional Sites (PROSITE) are used for identifying protein-coding genes based on RNA homology search.^[6] Previously, a number of statistical models have been proposed to achieve this goal. For example, covariance models (CMs),^[7] Pair Stochastic Tree Adjoining Grammars (PSTAG),^[8] Stochastic Context-free Grammars (SCFGs),^[9] and Iterated Loop Matching (ILM)^[10] have algorithms with high computational complexity and, therefore, increase the time needed to perform an alignment. Prediction of RNA structure is invaluable in creating new drugs and understanding genetic diseases. Several deterministic algorithms and soft computing-based techniques have been developed for more than a decade to determine the structure from a known RNA sequence.^[11] Thus, we have tried to provide models that speed up this process. One of the basic models for creating a test sequence search for an RNA family is hidden Markov model (HMM). In this way, the possibility of different position of symbols can be effectively described. The proposed model in this article, based on the context-sensitive HMM (csHMM) that uses this algorithm, can be provided a framework for the alignment of RNA secondary structures including pseudoknots.^[12]

Like protein-coding genes, ncRNA sequences can be gathered into groups of related sequences. We observed

sequences that have a spot with the same family and perform comparable functions (or functions that are associated in certain ways) in the cellular mechanism. In general, singular sequences in the family share one or more basic measurable components with different sequences that have a place with the same family. Such sequences are said to be homologous to each other, and henceforth called homologous. Given another sequence, we can exploit these family-particular attributes to figure out if it has a place with a particular sequence family. Its membership in a specific family can regularly be utilized to deduce the function of the sequence. The sequence-based methods (BLAST, FASTA, PROSITE, profileHMM) are very useful for identifying homologous DNAs and proteins, but they often behave poorly when applied to RNA homology search. The main reason that many functional ncRNAs preserve their secondary structures as well as their primary sequences. Sometimes, these base paired structures are still preserved among related RNAs, even when their similarity in the primary sequence level can be hardly recognized. Therefore, when evaluating the similarity between two RNA molecules, it is important to take both their primary sequences and their secondary structures into consideration. In many organisms, ncRNA genes do not show solid succession organization inclinations, which is the motivation behind why the customary methodologies that are principally taking into account base composition insights come up short. In this article, we propose another strategy that can be utilized for displaying and foreseeing RNA secondary structures. The proposed strategy depends on csHMMs. The article is organized as follows: in materials and methods section, we give a brief survey of RNA secondary structures and factual model for speaking to and investigating RNAs and survey the idea of csHMMs and clarify how csHMMs can be utilized as a part of RNA similitude seek. We then exhibit the execution of the proposed technique in results and discussion section and end with conclusion section.

Methods

We designed a suitable model for the alignment of secondary structures of RNAs and the proposed model depends on csHMMs. We tested our model using a few pseudoknots from the RNA family database (Rfam).^[13] The Rfam is a substantial accumulation of different RNA families, in which the member sequences in each family are adjusted to suit one another. In our trials, we used these sequences to achieve seed alignment in each RNA family as they were hand curated and have sensibly dependable structural annotation. For each sequence family, we picked and used one member as a reference RNA and its structural annotation to predict the secondary structure of other sequences in the same family.

Then, we tallied the quantity of accurately anticipated base pairs (true positives [TPs]), that of mistakenly anticipated

base pairs (false positives [FPs]), and that of base pairs in the classified structure that were not anticipated by the model (false negatives [FNs]). These numbers were used to compute the sensitivity (SN) and specificity (SP) of the system as follows:

$$SN = TP/(TP + FN) \quad SP = TP/(TP + FP) \quad (1)$$

RNA secondary structures

The second structure represents the linked pairs that arise after the folding of the RNA strand. Each nucleotide max can bond with other nucleotides. This article assumes that only the intended focal bonds can be formed, a pair of Watson–Crick (CG and AU and its inverse of the GC and UA), and the pair reversed its weak GU and UG because these pairs are more likely to occur and are abundant in the RNA molecule. Each combination of focal junctions denotes a secondary structure. In the case of the following three conditions for both the junctions $[i, j]$ and $[k, l]$ in which $i < j$ and $k < l$, the following three conditions are met:^[14]

- Each nucleotide can only bind to another nucleotide:
 $i = k$ if $j = l$
- Each nucleotide cannot bind to its side three nucleotides:
 $j - i > 4$
- Pseudoknots are not allowed, i.e., if we have $i < k < j$, then probably it should be $i < k < l < j$.

According to the rules that were developed for linking nucleotides, we can build all the links that could potentially be listed; in fact, The main issue in predicting the secondary structure of RNA is detected that which link is present in the secondary structure and which link does not present. An important feature is that the secondary structure of the molecule is stable in this case, that is, the links in the second structure appear that can match the shape of the molecule to equilibrium and stability. Generally, secondary structures of RNA are usually formed by one of the two hairpin loops or pseudoknot.^[3]

New studies on different genomes disclose the presence of many non-coding RNAs that, although not translated into proteins, act directly as RNAs and play an important role in different biological behaviors.^[6] Recent developments have shown that ncRNAs play an important role in silencing genes, processing and correcting RNAs, controlling transcription and translation, and many other regulatory actions.

RNA sequences often undergo compensatory mutations in order to preserve their secondary structures. For a given base pair in an RNA molecule, if the base in one side is changed to another base, the base in the other side is also changed such that the base pair is still maintained. As a result, we can observe strong correlation between the two base positions in homologous RNAs.^[15]

RNAs with secondary structure include one or more symmetrical area (or reverse complement area) in its first structure, and the pair of bases complementary

to the RNA folding is dumped. RNA sequences with secondary structures can be viewed as a kind of biological palindromes. Because of this symmetry, there is a strong correlation between the distant symbols. HMM, which can be considered as statistically regular grammars, cannot be used to describe the language of palindromes.

So far, the language of palindrome modeling of the higher-order grammars such as context-free grammars has been used. This grammars can describe the correlations between nesting symbols. SCFGs, as an example of this grammar, is widely used in the analysis of RNA sequences, is incapable of describing the correlation between crossing base pairs, and cannot sort of pseudoknot. In this article, to overcome this problem, instead of using grammars, csHMMs are used. These grammars can classify a large number of known pseudoknots.^[6]

Context-sensitive hidden Markov Models

Context-sensitive hidden Markov Model development of conventional hidden Markov Model

The csHMM is an extension of the traditional HMM.

In this model, there are different transition and emission probability states that are context dependent. Symbols that are emitted at certain states are stored in the memory, and the stored data serves as the context that affects the emission probabilities and the transition probabilities at certain future states. This context-sensitive property increases the descriptive power of the model significantly, compared to the traditional HMM.^[6]

It is assumed that csHMMs, including M states, are separate. The set of hidden state V is defined as follows:

$$V = S \cup P \cup C \cup \{\text{start, end}\} \quad (2)$$

According to Eq. 2, three different classes of states include single-emission states (S_n), pairwise-emission states (P_n), and context-sensitive states (C_n). S is the set of single-emission states, P is the set of pairwise-emission states, and C is the set of context-sensitive states, which are described as follows:

$$S = \{S_1, S_2, S_{M2}\} \quad (3)$$

$$P = \{P_1, P_2, P_{M1}\}, \quad C = \{C_1, C_2, C_{M1}\} \quad (4)$$

P_n and C_n states are always represented in pairs. For example, if we have two states, i.e., P_1 and P_2 , then HMM will require two states of C_1 and C_2 . Figure 1 shows an example of the dependence of P_n and C_n by a stack memory (Z_n). The process of observation for $X = x_1, x_2, \dots, x_L$ is determined so that x_i symbol in time i is observed. Each symbol x_i is accounted for a bit of alphabet $x_i \in A$, C , G , U (T).

Initially defined the probability that the model will make a transition from a state $s_i = v$ to the next state $s_{i+1} = w$. For $v \in S \cup P$, this probability defined as follow:

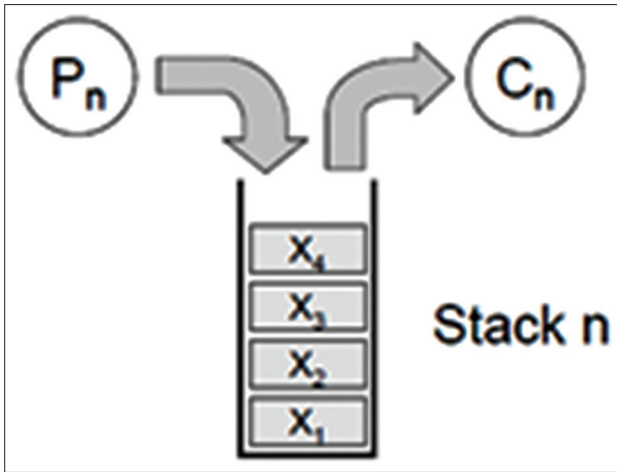


Figure 1: The states P_n and C_n associated with a stack Z_n

$$P(s_{i+1} = w | s_i = v) = t(v, w) \tag{5}$$

The transition probability state of a context-sensitive state is dependent on Z_n . The transition probability of “ $s_i = v \in C$ ” to “ $s_{i+1} = w$ ” for the sample can be as follows:

$$P(s_{i+1} = w | s_i = v, z_n) = t(v, w) \tag{6}$$

The emission probability is the probability of the observation symbol $x_i = x$ from the hidden state $s_i = v$. F or $v \in S \cup P$; the probability to be defined is as follows:

$$P(x_i = x | s_i = v) = e(x | v) \tag{7}$$

Because $\forall v \in C$ emission probability is dependent on $s_i = v$ and a symbol x_p in Z_n memory, it is defined as follows:

$$P(x_i = x | s_i = v, Z_n) = e(x | v, x_p) \tag{8}$$

Using csHMM, we can easily build a simple model that generates only palindrome. For example, the structure shown in Figure 2 can be used. As shown in the figure, we have three hidden states, namely S_1, P_1 and C_1 , so that pairing state (C_1, P_1) can be linked with a stack. Initially, the modeling starts at pairwise-emission state. This state will do several transfers to generate a number of symbols that go into the stack. When model enters the C_1 , related transition and emission probabilities are set, so that the symbols are always on the top of the stack in this state, are emitted, and the transfer continues until the stack is empty. In this way, C_1 emits the same symbols that are emitted from P_1 inversely.^[4]

If we consider that the number of symbols emitted from the stack is N , the generated sequences will always be a palindrome of the form $x_1 \dots x_N x_N \dots x_1$ (even-length sequence) or $x_1 \dots x_N x_{N+1} x_N \dots x_1$ (odd-length sequence).

Finding the most probable path

They are assumed to have the observation sequence of $X = x_1 x_2 \dots x_L$. As noted previously, if denote the underlying state of x_i as s_i . Assuming that there are M distinct states in the model, we have M^L different paths.

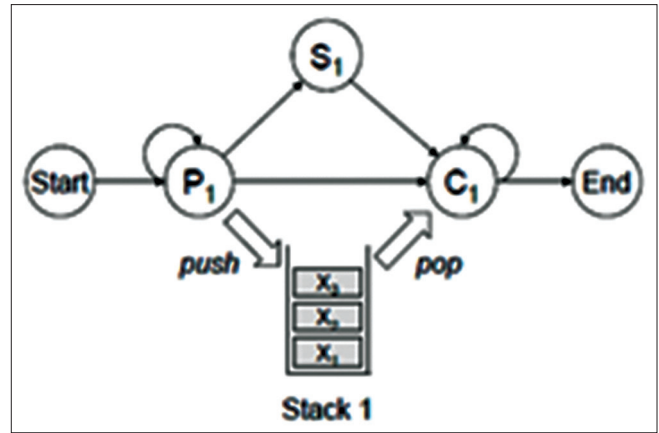


Figure 2: A case of context-sensitive hidden Markov model that generates only palindromes

Given the observation sequence x , how can we find the path that is most probable among the M^L distinct paths?

Thus, in this article, alignment algorithm was designed using a csHMM to calculate the log-like probability of optimal path. In $v \in P \cup C$, \bar{v} is defined to supplement v as follows:

$$v = P_n \rightarrow \bar{v} = C_n, v = C_n \rightarrow v = P_n \tag{9}$$

In $v \in P \cup S$, the probability of obtaining x from v is already defined for $e(x | v)$ and that of $v \in C$ is already defined for $e(x | v, x_p)$, where x_p represents the emission from pair-wise state (\bar{v}). The transition probability from state v to w is defined by $t(v, w)$. Finally, $\gamma(i, j, v, w)$ defines the log-like probability of the optimal path among all the s_i, s_j subpaths, where $s_i = v$ and $s_j = w$.

Estimating the model parameters

To apply the csHMM to real-world problems, adjusting the model parameters is critical to optimize the method. Thus, finding a way to optimize parameter θ is essential in order to maximize the probability $P(x \cup \theta)$ obtained from x . The process of finding these parameters is often called training; however, finding an analytical solution for model parameters is impractical. Thus, in this article, expectation-maximization (EM) algorithm was used to obtain the local maximum $P(x \cup \theta)$.

Results

To test the alignment algorithm, the hepatitis delta virus (HDV) ribozyme was first checked. One of the family’s members was chosen randomly to train the data. The csHMM used an algorithm to identify the RNA as shown in Figure 3. The model in the figure has 14 single-emission states s_1, s_2, \dots, s_{14} and 6 sets of pairwise-emission and context-sensitive states. Each pair, that is, $(p_1, c_1), (p_2, c_2)$, and (p_6, c_6) , is connected with a different stack.

Using an alignment algorithm, the optimal path for this RNA is defined as follows:

$$S^* = s_1 p_1 s_2 p_2 s_3 c_2 p_2 s_4 p_3 s_5 p_4 p_4 s_6 p_5 s_7 c_5 s_8 c_4 s_9 p_6 s_{10} c_6 s_{11} c_4 s_{12} c_3 s_{13} c_2 c_1 s_{14} \quad (10)$$

Table 1 shows the emission probabilities which were estimated using the EM algorithm after ten iterations.

By comparing Tables 1 and 2, it can be seen that the estimated values were close to the original values.

In purine riboswitches, the RNA acts as an HDV ribozyme. Again, the csHMM used an algorithm to identify the RNA, as shown in Figure 4.

The model shown in the figure has 11 single-emission states, that is, s_1, s_2, \dots, s_{11} , and 5 pairs of pairwise-emission and context-sensitive states. Each pair, that is, $(p_1, c_1), (p_2, c_2), \dots,$ and (p_5, c_5) , is associated with a separate stack. The emission probabilities are shown in Table 3.

As with HDV ribozyme, an alignment algorithm was used to define the optimal path for this RNA as follows:

$$S^* = s_1 p_1 s_2 p_2 s_3 p_3 s_4 p_4 s_5 c_4 s_6 c_3 s_7 p_5 s_8 c_5 s_9 c_2 s_{10} c_1 s_{11} \quad (11)$$

Table 4 shows the emission probabilities which were estimated using the EM algorithm after ten iterations.

By comparing Tables 3 and 4, it can again be observed that the estimated values were close to the original values.

Discussion

To acquire solid evaluations of the above results, we performed a cross-validation test by repeating the same procedure for fifty members in each given RNA family. Then, we registered the general prediction ratios in view of the log-like probability values of the optimal path obtained. The accuracy of the

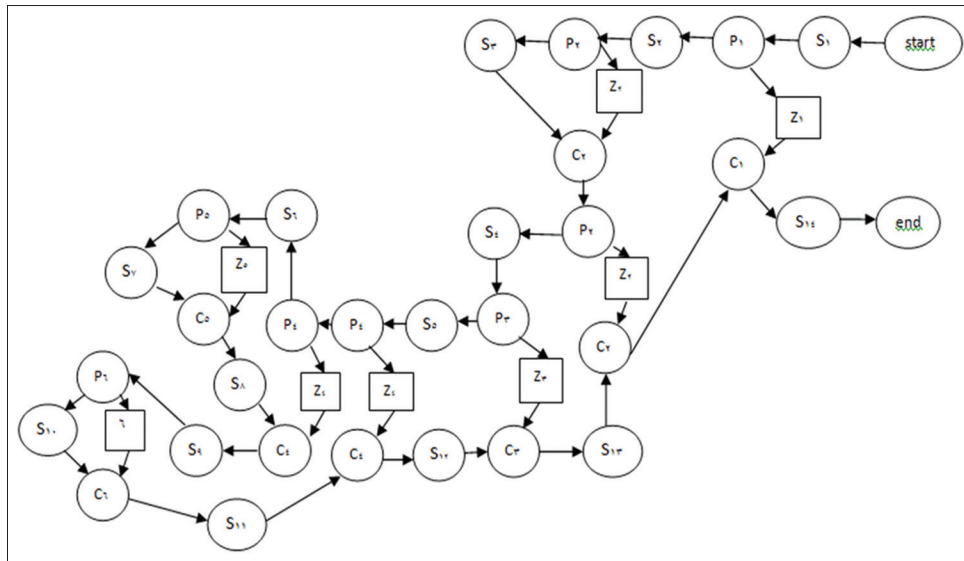


Figure 3: Hepatitis delta virus ribozyme model using context-sensitive hidden Markov model

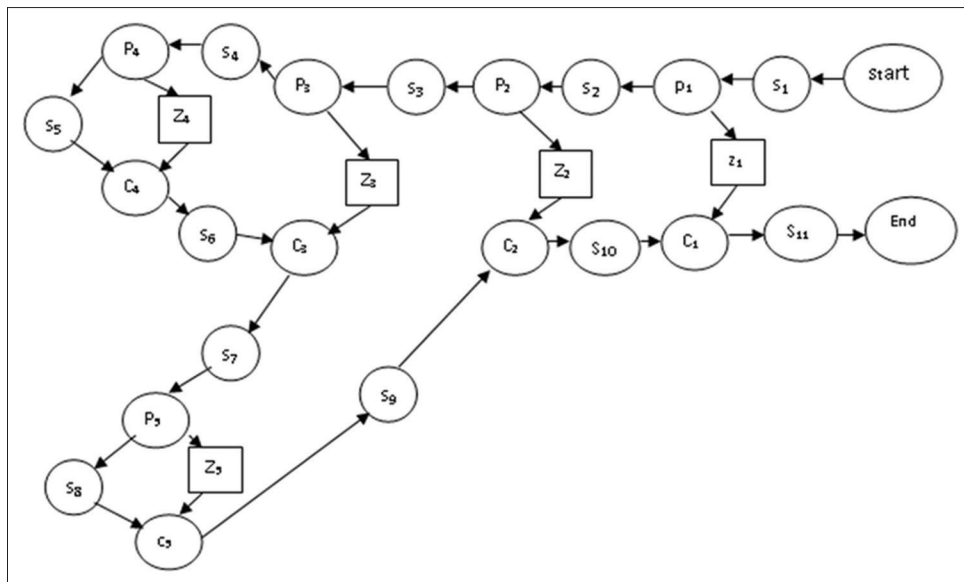


Figure 4: Purine riboswitch model using context-sensitive hidden Markov models

Table 1: Estimated emission probabilities $e(x|v)$ for hepatitis delta virus ribozyme

	A	C	G	U
P_1	0.0088	0.1818	0.8094	0.0000
S_1	0.0017	0.2715	0.3983	0.3285
P_2	0.4067	0.0628	0.0400	0.4905
S_2	0.9983	0.0007	0.0000	0.0010
S_3	0.1664	0.5006	0.0036	0.3294

A – Adenine; C – Cytosine; G – Guanine; U – Uracil

Table 2: Emission probabilities $e(x|v)$ for hepatitis delta virus ribozyme

	A	C	G	U
P_1	0.000	0.250	0.750	0.000
S_1	0.000	0.285	0.430	0.285
P_2	0.000	0.444	0.222	0.333
S_2	1.000	0.000	0.000	0.000
S_3	0.200	0.400	0.000	0.400

A – Adenine; C – Cytosine; G – Guanine; U – Uracil

Table 3: Emission probabilities $e(x|v)$ for purine riboswitch

	A	C	G	U
P_1	0.10	0.15	0.45	0.30
S_1	0.13	0.34	0.07	0.46
P_2	0.39	0.06	0.08	0.47
S_2	0.25	0.40	0.21	0.14
S_3	0.08	0.12	0.73	0.07

A – Adenine; C – Cytosine; G – Guanine; U – Uracil

model in determining the members of the HDV family was 83.33%, whereas the accuracy of the model in determining the members of the purine family was 65%.

To compare the performance of the proposed method with that of PSTAGs and that of ILM, we tested the accuracy of the algorithm across the HDV RNA family, which has a pseudoknot structure. Table 5 compares the prediction results of the proposed strategy with those of PSTAGs and ILM. As shown in the table, csHMM yielded accurate prediction results that were practically identical to those of PSTAGs and ILM.

In addition, we tested the performance of the proposed method in the purine family, which has a more complicated secondary structure than the past RNA family; thus, PSTAGs nor ILM can take care of the purine family. Table 6 demonstrates the prediction results of the proposed method with those of a profile-HMM; it shows that the csHMM delivered preferable alignments over the conventional profile-HMM.

csHMMs can reflect the pair-wise interactions between distant symbols adequately to model and predict RNAs that conserve secondary structures as well as primary sequences. In addition, csHMMs can be used to model pseudoknots, in which the

Table 4: Estimated emission probabilities $e(x|v)$ for purine riboswitch

	A	C	G	U
P_1	0.1510	0.1622	0.3431	0.3437
S_1	0.1439	0.3041	0.0911	0.4609
P_2	0.4142	0.0513	0.0717	0.4628
S_2	0.4012	0.3745	0.2078	0.0165
S_3	0.0750	0.1064	0.7925	0.0261

A – Adenine; C – Cytosine; G – Guanine; U – Uracil

Table 5: Performance comparison between csHMM, ILM and PSTAG

csHMM (%)		ILM (%)		PSTAG (%)	
SN	SP	SN	SP	SN	SP
97	89	85.70	82.40	96	88.90

csHMM – Context-sensitive hidden Markov Model; ILM – Iterated Loop Matching; PSTAG – Pair Stochastic Tree Adjoining Grammar; SN – Sensitivity; SP – Specificity

Table 6: Performance comparison between csHMM and Profile-HMM

csHMM (%)		Profile-HMM (%)	
SN	SP	SN	SP
76	76	58	65

HMM – Hidden Markov Model; csHMM – Context-sensitive hidden Markov Model; SN – Sensitivity; SP – Specificity

pair-wise interactions between bases are permitted to cross one another. Pseudoknots are found in numerous RNAs, and distinguishing them is essential in some applications, for example, in three-dimensional structure predictions.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

- Miranda-Rios J, Navarro M, Soberón M. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc Natl Acad Sci U S A* 2001;98:9736-41.
- Moulton V. Tracking down noncoding RNAs. *Proc Natl Acad Sci U S A* 2005;102:2269-70.
- Fiannaca A, La Rosa M, La Paglia L, Rizzo R, Urso A. NRC: Non-coding RNA classifier based on structural features. *BioData Min* 2017;10:27.
- Schneider HW, Raiol T, Brigido MM, Walter ME, Stadler PF. A support vector machine based method to distinguish long non-coding RNAs from protein coding transcripts. *BMC Genomics* 2017;18:804.
- Byung JY, Vaidyanathan PP. Structural alignment of RNAs using profile-csHMMs and its application to RNA homology search: Overview and new results. *Automat Control IEEE Trans* 2008;53:10-25.
- Byung JY. *Signal Processing Methods for Genomic Sequence*

- Analysis. Thesis Presented to the California Institute of Technology for the Degree of Doctor of Philosophy. Pasadena, California; 2007.
7. Eddy SR, Durbin R. RNA sequence analysis using covariance models. *Nucleic Acids Res* 1994;22:2079-88.
 8. Hiroshi M, Kengo S, Yasubumi S. Pair stochastic tree adjoining grammars for aligning and predicting pseudoknot RNA structure. *Bioinformatics* 2005;21:2611-7.
 9. Lari K, Young SJ. The estimation of stochastic context-free grammars using the inside-outside algorithm. *Comput Speech Lang* 1990;4:35-56.
 10. Ruan J, Stormo GD, Zhang W. An iterated loop matching approach to the prediction of RNA secondary structures with pseudoknots. *Bioinformatics* 2004;20:58-66.
 11. Shubhra SR, Sankar KP. RNA secondary structure prediction using soft computing. *IEEE/ACM Trans Comput Biol Bioinform* 2013;10.
 12. Christopher N. Hidden Markov Models with Applications to DNA Sequence Analysis, *STOR-i*; May, 2011.
 13. Available from: <http://rfam.sanger.ac.uk/>; <http://rfam.janelia.org/>. [Last accessed on 2015 Jul 14; Last accessed on 2019 Jun 22].
 14. Richard D, Sean E, Anders K, Graeme M. *Biological Sequence Analysis*. Cambridge, UK: Cambridge University Press; 1998.
 15. Byung JY. Comparative analysis of biological networks: Hidden Markov model and Markov chain-based approach. *Signal Process Mag IEEE* 2012;29:22-34.

BIOGRAPHIES



Nayyer Mostaghim Bakhshayesh was born in Tabriz, Iran, in 1987. She graduated from Farzaneghan high school (major: Mathematics-Physics) in Tabriz, Iran, in 2005. She received the B.Sc. degree in Electrical Engineering (major: Biomedical Engineering) in 2010 and M.Sc. degree in Biomedical Engineering in 2015, all from Sahand University of Technology, Tabriz, Iran. She is currently a PhD student in Biomedical Engineering from Sep. 2016 at Sahand University of Technology, Tabriz, Iran.

Her research interests include Bioinformatics, Genomic Signal Processing and Microarray Image Processing.

Email: n_mostaghim@sut.ac.ir



Mousa Shamsi received his B.Sc. degree in Electronic Engineering from Tabriz University, Tabriz, Iran, in 1995. He received his M.Sc. degree in Biomedical Engineering from University of Tehran, Tehran, Iran, in 1999. From 1999 to 2002, he taught as a lecturer at the Sahand University of Technology, Tabriz, Iran.

From 2002 to 2008, he was a PhD student at the University of Tehran, Tehran, Iran, in Bioelectrical Engineering. In 2006, he was granted with the Iranian government scholarship as a visiting researcher at the Ryukyus University, Okinawa, Japan. From December 2006 to May 2008, he was a visiting researcher at this University. He received his PhD degree in Biomedical Engineering from University of Tehran, Tehran, Iran, in December 2008. From December 2008 to April 2013, he was an assistant professor at Faculty of Electrical Engineering, Sahand University of technology, Tabriz, Iran.

From April 2013, he is an Associate Professor at Faculty of Biomedical Engineering, Sahand University of technology, Tabriz, Iran. His research interests include Biomedical Signal and Image Processing, Pattern Recognition and Machine Learning.

Email: shamsi@sut.ac.ir



Mohammad Hossein Sedaaghi was born in Tehran, Iran. He received the B.Sc. and M.Sc. degrees from the Sharif University of Technology, Tehran, IRAN, in 1986 and 1987, respectively. In 1998, he received the Ph.D. degree from Liverpool University. He is now a professor at Sahand University of Technology, Tabriz, Iran. His research interests include signal/image processing, pattern recognition, machine learning and biometrics.

Email: sedaaghi@sut.ac.ir



Hossein Ebrahimnezhad received his BS and MS in electronic and communication engineering from Tabriz University, Iran and K.N.Toosi University of Technology, Iran in 1994 and 1996, respectively. In 2007, he received his PhD from Tarbiat Modares University, Iran. Currently, he is a professor at Sahand University of Technology, Tabriz,

Iran. His research interests include image and multimedia processing, computer vision, 3D model processing and soft computing.

Email: ebrahimnezhad@sut.ac.ir