Evolutionary Layered Hypernetworks for Identifying microRNA-mRNA Regulatory Modules

Soo-Jin Kim*, Jung-Woo Ha*, Bado Lee and Byoung-Tak Zhang

Abstract— **Exploring microRNA (miRNA) and mRNA regulatory interactions may give new insights into diverse biological phenomena. While elucidating complex miRNA-mRNA interactions has been studied with experimental and computational approaches, it is still difficult to infer miRNA-mRNA regulatory modules. Here we present a novel method for identifying functional miRNA-mRNA modules from heterogeneous expression data. The proposed approach is layered hypernetworks consisting of two layers which are the layer of modality-dependent hypernetworks and of an integrating hypernetwork. The layered hypernetwork model is suitable for detecting relationships between heterogeneous modalities. Applied to the analysis of miRNA and mRNA expression profiles on multiple human cancers, the proposed model identifies oncogenic miRNA-mRNA regulatory modules. The experimental results show that our method provides a competitive performance to support vector machines, and outperforms other standard machine learning algorithms. The biological significance of the discovered miRNA-mRNA modules were validated by literature reviews.**

I. Introduction

RECENTLY, microRNAs (miRNAs) have been discovered as important regulators that play a major role in various cellular processes, such as cell differentiation, proliferation, growth, mobility and apoptosis. Multiple studies involving certain types of cancers have proved that miRNAs have a crucial role in tumor progression by regulating target mRNAs. Therefore, it is essential to identify functional interactions between miRNA and mRNA for understanding the context-dependent activities of miRNAs in complex biological systems. However, the precise regulatory mechanisms of miRNAs and mRNAs remain to be elucidated.

Early work in this area primarily focused on the genome-wide computational prediction of miRNAs [1] and their targets [2]. The miRNAs and mRNAs participate as important components of gene regulatory networks in diverse biological processes together. The exploration of miRNA and mRNA interactions can be an essential first step toward the discovery of their combinational effects on different physiological conditions. To address such issues, several efforts have been made to detect miRNA-mRNA regulatory relationships. Yoon *et al.* [3] proposed a computational method to predict the groups of miRNAs and mRNAs using weighted bipartite graph models. But, predictions only based on sequence information may be insufficient to determine the complex regulatory modules of miRNA and mRNA. Huang *et al*. [4] provided a probabilistic framework which applied Bayesian network parameter learning to detect miRNA-mRNA interactions. Also, Joung *et al.* [5] used a bi-clustering approach based on co-evolutionary learning strategies to discover miRNA-mRNA modules. Their methods integrate multiple sources to identify miRNA-mRNA pairs. Because of the complexity and diversity of miRNA and mRNA interactions, however, inferring functional miRNA-mRNA regulatory modules remains a difficult problem.

In this paper, we propose a novel approach for identifying miRNA-mRNA regulatory modules associated with cancers from expression data. The proposed method is a new modified hypernetworks, so called layered hypernetworks (LHNs). A hypernetwork is a generalized hypergraph which can represent higher-order relationships among vertices [6]. The LHN is an advanced hypernetwork model having a hierarchical structure where evolutionary computation is embedded in the learning process. Given data consisting of more than one modality, an LHN is composed of two layers. The first layer's hypernetworks are built from each modality of data. And the second layer's hypernetwork represents intermodal relationships by combining the learning results of the first layer's hypernetworks. This property is useful for analyzing complicated and heterogeneous problems such as the identification of miRNA and mRNA interactions. In the LHN frameworks, learning process is performed by an evolutionary algorithm to find the best combinations of higher-order elements in a huge combinatorial searching space.

In experiments, we apply the LHN model to miRNA and mRNA expression profiles related to multiple human cancer [7], [8]. The goal is to identify functional miRNA-mRNA regulatory interactions with high accuracy. Our experimental results show that the proposed method provides a competitive performance to support vector machines and outperforms decision trees, Bayesian networks, naïve Bayes and random

This work was supported in part by IT R&D Program of MKE/KEIT (KI002138, MARS), in part by NRF Grant of MEST (314-2008-1-D00377, Xtran), and in part by the BK21-IT program funded by Korean Government (MEST). The ICT at Seoul National University provided research facilities for this study.

B.-T. Zhang is with the Center for Biointelligence Technology, the School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea (corresponding author to provide phone: +82-2-880-1833; fax: +82-2-875-2240; e-mail: btzhang@bi.snu.ac.kr).

S.-J. Kim is with Interdisciplinary Program in Bioinformatics, the Center for Biointelligence Technology, Seoul National University, Seoul 151-742, Korea (email: sjkim@bi.snu.ac.kr).

J.-W. Ha is with the School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea (e-mail: jwha@bi.snu.ac.kr).

B. Lee is with the School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea (e-mail: bdlee@bi.snu.ac.kr).

*: Both authors have equally contributed to this work.

forests. We also validated the biological significance of the discovered miRNA-mRNA modules by literature search.

The paper is organized as follows. In Section 2, the conventional hypernetwork models are explained. Section 3 describes our proposed approach, layered hypernetwork models. Section 4 elucidates the evolutionary learning procedure for LHNs. In Section 5, the experimental results on miRNA and mRNA expression profiles are provided. Concluding remarks and directions for further works are given in Section 6.

II. HYPERNETWORK MODELS

A hypernetwork is a weighted random hypergraph model inspired by biomolecular networks [6]. Here we briefly introduce the hypernetwork models. In hypernetworks, a vertex denotes a value of attributes and a hyperedge represents an arbitrary higher-order combination of vertices with its own weight. Formally, a hypernetwork H is defined as $H = (V, E, W)$, where V, E and W are a set of vertices, hyperedges, and weights respectively. A hyperedge of order (cardinality) k is referred to a k-hyperedge. A hypernetwork is called a k-hypernetwork where all hyperedges are k-hyperedges. Fig. 1 shows an example of the hypernetwork with binary data consisting of eight attributes. Since a hypernetwork can be considered as a probabilistic memory model to store segments of information of data, $D = \{\mathbf{x}^{(n)}\}_{n=1}^{N}$, it can be formulated with the energy function of stored data and parameters. The energy of the k-hypernetwork is defined as follow:

$$\varepsilon(\mathbf{x}^{(n)}; W) = -\sum_{i=1}^{|E|} w_i^{(k)} I(\mathbf{x}^{(n)}, E_i), \quad (1)$$

where $w_i^{(k)}$ is a weight of i-th hyperedge E_i with k-order, $\mathbf{x}^{(n)}$ is the n-th example pattern to store and $I(\mathbf{x}^{(n)}, E_i)$ denotes the combination of elements of $\mathbf{x}^{(n)}$ in hyperedge E_i. Then, the probability of the data being generated from the hypernetwork is given as a Gibbs distribution:

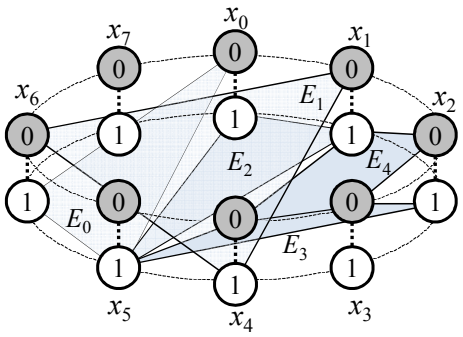
$$P(\mathbf{x}^{(n)} \mid W) = \frac{1}{Z(W)} \exp\left\{-\varepsilon(\mathbf{x}^{(n)}; W)\right\}, \quad (2)$$

where $Z(W)$ is a partition function. The partition function $Z(W)$ can be formulated as:

$$Z(W) = \sum_{\mathbf{x}^{(m)} \in D} \exp\left\{-\varepsilon(\mathbf{x}^{(m)}; W)\right\} = \sum_{\mathbf{x}^{(m)} \in D} \exp\left\{\sum_{i=1}^{|E|} w_i^{(k)} I(\mathbf{x}^{(m)}, E_i)\right\}. \quad (3)$$

That is, a hypernetwork is described with a probability distribution of combinations of random variables with weights as parameters. A likelihood function also is maximized by finding hyperedge compositions which can reveal the distribution of given data better [6], [11].

For evolving hypernetworks, we assume that a population is a hypernetwork and its individuals are hyperedges. A change of a set of hyperedges leads to evolve a structure of hypernetworks. Hyperedge compositions (hypernetwork structure) and hyperedge weights (parameters) are learned by evolutionary processes using the general operations of matching, selection, amplification, deletion of hyperedges. So the population is converged to the optimal hypernetworks that maximize the performance by evolutionary self-organizing process. References [9]-[12] provide various evolutionary strategies in the learning process of hypernetworks. Note that the learning task of a hypernetwork is focused on evolving the whole structure of a hypernetwork (population) rather than the respective hyperedges (individuals).



Edges	Vertices	Order
E_0	$\{x_0=0, x_5=1, x_6=1, W_0\}$	3
E_1	$\{x_1=1, x_4=1, x_6=0, W_1\}$	3
E_2	$\{x_0=1, x_1=1, x_5=1, W_2\}$	3
E_3	$\{x_2=1, x_3=0, x_4=0, x_5=1, W_3\}$	4
E_4	$\{x_1=1, x_2=0, x_3=0, x_4=0, W_4\}$	4

Fig. 1. An example of a hypernetwork with binary data consisting of eight attributes and corresponding table of hyperedges.

III. LAYERED HYPERNETWORK MODELS

An LHN is an advanced hypernetwork model with hierarchical structures consisting of two layers. The first layer is the layer of modality-dependent hypernetworks which reflects the property of data consisting of more than one modality. The second layer is an integrating layer to analyze relationships between each modality from the first layer's hypernetworks. In the first layer, the same number of hypernetworks exists as the number of modalities of the given data, and each hypernetwork is built from samples of a modality. Dissimilar to the first layer, only one hypernetwork exists in the second layer. Hyperedges of the second layer's hypernetwork are generated by combining randomly selected hyperedges from each modality hypernetwork. Therefore the hypernetwork of the second layer represents relationships among several modalities. Same as conventional hypernetworks, formally, the second layer's hypernetwork is defined with the energy function when provided a weight vector as a parameter. When given a dataset D with two modalities, $D = \{\mathbf{x}^{(n)}\}_{n=1}^{N} = \{(\mathbf{m}^1, \mathbf{m}^2)^{(n)}\}_{n=1}^{N}$, the energy of the second

layer's hypernetwork $\varepsilon(\mathbf{x}^{(n)};W)$ generated from the first layer's *k*-hypernetworks is defined as follows:

$$\varepsilon(\mathbf{x}^{(n)};W) = \varepsilon\left\{(\mathbf{m}^1,\mathbf{m}^2)^{(n)};W\right\} = -\sum_{i=1}^{|E|} w_i^{(k)} I\left\{(\mathbf{m}^1,\mathbf{m}^2)^{(n)},E_i\right\}, \quad (4)$$

where $\mathbf{m}^1$ and $\mathbf{m}^2$ are vectors of each modality variable which make up of a data sample $\mathbf{x}$. Same as (2), then the probability of generating *n*-th data with two modalities, $P(\mathbf{x}^{(n)}|W)$ is expressed as:

$$P(\mathbf{x}^{(n)}\mid W) = \frac{1}{Z(W)}\exp\left[-\varepsilon\left\{(\mathbf{m}^1,\mathbf{m}^2)^{(n)};W\right\}\right]. \quad (5)$$

Fig. 2 shows the architecture of a LHN for miRNA and mRNA expression data.

IV. Evolutionary Learning of Layered Hypernetworks

A. Evolving Hypernetworks in the First Layer

Evolving process of hypernetworks in the first layer consists of three parts: building, learning, and evaluating like conventional hypernetworks. Same as the general evolutionary methods, it is important to decide the size of population in the building procedure of hypernetworks. The reason is that the size of a hypernetwork determines the coverage of underlying patterns of data as well as the diversity of the population. That is, it is needed that the size of a hypernetwork is larger than a specific threshold for a hypernetwork without any omitted attribute values. So, for generating a *k*-hypernetwork which misses *i* random variables, let us denote the number of omitted variables in a *k*-hypernetwork. Then, Pr($x = i$) is calculated as follows:

$$\Pr(x=i) = {}_K C_i \left(\frac{{}_{K-i}P_k}{{}_K P_k}\right)^{N\times r} = {}_K C_i \left(\frac{(K-i)!(K-k)!}{K!(K-i-k)!}\right)^{N\times r}, \quad (6)$$

where *K* is the number of attributes in data, *k* is the order of hyperedges, *N* is the size of data, and *r* is a sampling rate which means the number of generated hyperedges from an example. That is, $N \times r$ is equal to the number of hyperedges in a hypernetwork as a population size. Therefore, the probability that no omitted variables exists in a hypernetwork, Pr(x = 0) is

$$\Pr(x=0) = 1-\sum_{i=1}^{K-k}\Pr(x=i) = 1-\sum_{i=1}^{K-k} {}_K C_i \left(\frac{(K-i)!(K-k)!}{K!(K-i-k)!}\right)^{N\times r}. \quad (7)$$

According to (7), assuming that the size of data is fixed, the sampling rate *r* and the order of hyperedges *k* determine whether omitted variables exist or not. That is, if *r* and *k* get larger, Pr(x = 0) decreases exponentially. Fig. 3 shows the changing pattern of omitted variables as the increasing of a population size and the order of hyperedges. Since higher-order hyperedges can represent only a specific pattern and are less likely to match examples in the learning process, however, too large *k* can lead a decrease of the classification accuracy. Also, too large population size causes a high cost to evolve hypernetworks. Therefore it is important to determine a suitable order and sampling rate to obtain good classification accuracy and to perform the learning process within reasonable time. Learning of each hypernetwork is based on comparing values of vertices in a hyperedge with their corres-

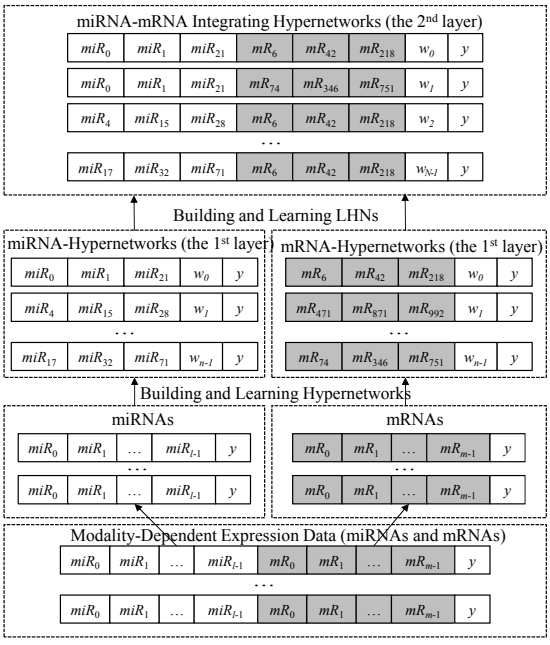


Fig 2. Architecture of LHN for miRNA and mRNA expression data. The LHN model is made up of two layers of hypernetworks. Hypernetworks in the first layer are constructed reflecting each modality of miRNA and mRNA data and the second layer's hypernetwork represents intermodality relationships between miRNA and miRNA hypernetworks of the first layer.

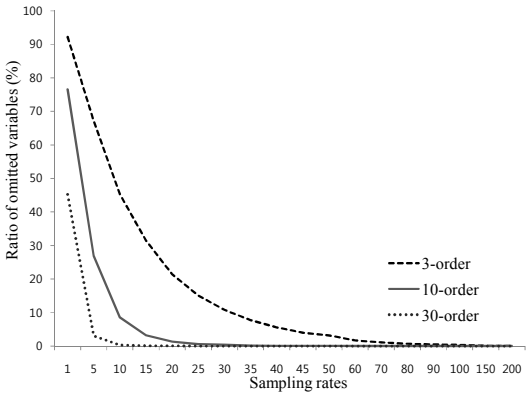


Fig 3. Omitted variables decrease as the population size increases. As sampling rates and the order of hyperedges are increasing, the ratio of omitted variables decreases exponentially.

ponding values in training examples [9]. The fitness of hyperedges (so called weight) is updated by the learning process. The weight is calculated based on the number of correct and incorrect prediction on training dataset. In this study we obtain the weight of hyperedges w as linear summation of two factors as follow. Where $\#_c$ and $\#_w$ are the number of correctly and incorrectly predicted training examples,

$$w = \alpha \times (\#_c - (\#_w)^2) + (1-\alpha) \times (\#_c + \#_w) \quad . \tag{8}$$

In (8), the first term is an accuracy factor because the score is higher when correctly predicted data are more and incorrectly ones are less. The second term reveals the coverage of hyperedges for data, which means that the second score is higher if hyperedges match more training examples. We introduced the replacement policy of hyperedges based on their weight as a strategy to enhancing the diversity of a hypernetwork in our previous study. However, the ratio of replaced hyperedges was fixed with a specific value [10]. In this study, we adopt the flexible ratio of hyperedge replacement to evolve hypernetworks of the first layer. The replacing ratio $R(t)$ is a function of the generation number in the evolution process and it is defined as follows:

$$R(t) = (M - m) \times \exp\left(\frac{1-t}{c}\right) + m \quad , \tag{9}$$

where M and m are the maximum and minimum value of replacing rates, which are constants in boundary of (0, 1), t means the generation of evolution, and c is an arbitrary positive constant. According to (9), in the first generation of evolution, $R(t)$ starts at M and a value converges to m as the evolution proceeds. In addition, we can control the converging speed to the minimum of a replacing ratio by the changing c.

B. Evolving a Hypernetwork in the Second Layer

A hypernetwork in the second layer represents the relationships between modalities. Since hyperedges in a hypernetwork of the second layer are generated by combining hyperedges of each hypernetwork in the first layer, the second layer's hypernetwork represents interactions between each other modality.

The maximum population size of the second layer's hypernetwork can be equal to the size of products of the first layer's hypernetworks. We integrate the hyperedges based on random selection to generate hyperedges of the second layer's hypernetwork, because it is not feasible to learning all combinations of hyperedges from the first layer's hypernetworks. Considering the causality of inter-modalities, for all hyperedges of the causal modality hypernetwork, the fixed size of randomly selected hyperedges are combined to generate the second layer's hypernetwork from the resultant modality hypernetworks. The fixed size of selected resultant hyperedges per causal hyperedges is called to the combining ratio.

H_i: hypernetworks of modality i in the first layer
LH: the second-layer's hypernetwork
M^i: separate data of modality i
$R(i)$: replacing ratio of i-th generation
CR: size of combined hyperedges of H_2 per hyperedge of H_1
y_i : class label of E_i
$H_1 \leftarrow$ Makehypernetwork(M^1);
$H_2 \leftarrow$ Makehypernetwork(M^2);
For $i \leftarrow 1$ to End condition
  $LH \leftarrow \{\}$;
  Evolvehypernetwork(H_1);
  Evolvehypernetwork(H_2);
  For $j \leftarrow 1$ to $|H_1|$
    $E_1 \leftarrow$ the j-th hyperedges of H_1
    For k $\leftarrow 1$ to CR
      $idx \leftarrow$ Random($|H_2|$);
      $E_2 \leftarrow$ the idx-th hyperedges of H_2
      If y_1 is equal to y_2
        $E \leftarrow E_1 \cup E_2$;
        $LH \leftarrow LH \cup E$;
      Otherwise re-select idx;
      End If
    End For
  End For
  EvaluateonTrainingData(LH);
  $H_1 \leftarrow$ RemoveAndResample(H_1, R(i));
  $H_2 \leftarrow$ RemoveAndResample(H_2, R(i));
End For
EvaluateonTestData(LH);

Fig 4. The algorithm of generating, evolving and learning LHNs. Used functions such as Makehypernetwork(.), Evolvehypernetwork(.), RemoveAndResampling(.), and Evaluate(.) are explained in our previous studies [9]-[11]. Generally learning finishes after the fixed amounts of epochs

The different point of LHNs from conventional hypernetwork models is that there exists crossover effect. In traditional hypernetwork models, a hyperedge is sampled from an example and there is not any crossover operation among hyperedges. However, by the process of separating and combining attributes with randomness, crossover operations are carried out to evolve the second layer's hypernetworks. Fig. 4 depicts the evolutionary algorithm of LHN models.

C. Significance Score Metric for hyperedges analysis

To find the significant relationships between modalities, we define a metric in this study. The proposed metric named significance score S is calculated based on co-occurrence frequency in the hyperedges of attributes of each modality. That is, the number that each pair of two attributes appears in all hyperedges becomes S. When data consists of two modalities, the proposed significant score S becomes a matrix whose size is l by m, where l and m are the size of modality 1 and modality 2 respectively. Significance score s_{ij} between the i-th attribute of a modality and the j-th attribute of the other modality, which is an element of S is defined to:

$$s_{ij} = \sum_{n=1}^{N} w_{ij}, w_{ij} = \begin{cases} 1 & ((i,j) \in E_n) \\ 0 & (otherwise) \end{cases} . \quad (10)$$

The score is the number of hyperedges which have the *i*-th and *j*-th attribute values simultaneously as their vertices. Therefore, a pair of attributes with high significance score can be regarded as a relevant interacting module.

V. Experimental Results

A. Data preparation and Experimental Setting

For experiments, we used miRNA and mRNA expression data associated to human epithelial cancers [7], [8]. The data have the expression profiles of 151 miRNAs and 10,262 mRNAs from samples consisting of 21 normal tissues and 68 multiple cancer tissues. We use a set of data (x, y), where x = $(x_1, x_2, \ldots, x_n) \in \{0, 1\}^n$ and $y \in \{0, 1\}$; i.e., a binary dataset. Although LHN models can process any attributes such as integers or real numbers, the discretized values simplify the representation of objectives and provide an efficient implementation of the LHN algorithm *in silico*. Hence, we preprocess the expression dataset by two different ways, gene-wise normalization and sample-wise binarization. We normalize the expression profiles based on the average of its gene values for each sample. And then, we convert the normalized expression values into binary numbers using the average of its samples for each gene. To obtain biologically meaningful results, also we extract a subset of genes from total mRNAs according to [13] and [14].

The experimental parameters are shown in Table 1. According to Section 4.A, we set up 80 as a sampling rate to avoid unrepresented variables with considering computational costs. Also, we determine α value for the accuracy term to have the influence on calculating weights. Considering regulatory mechanism of miRNA-mRNA, in addition, both miRNA and mRNA become a causal and resultant modality in turn.

TABLE 1. THE PARAMETERS USED FOR THE EXPERIMENT.

Parameters	Value
Order (# miRNA, # mRNA)	(3, 3)
Replacing rate (Min., Max.)	(0.1, 0.9)
# Sampling rate	80
# Combining rate	10
Ratio of accuracy score (α)	0.8
Num. of generations	50
Causal modality	both
Resultant modality	both

SR means a sampling rate, α is the ratio of accuracy score in obtaining weight of hyperedges, and CR represents the number of combination between hyperedges of miRNA and mRNA hypernetworks in the second layer's hypernetwork. In this study, miRNA and mRNA become causal modality in turn because both of them can be the causal factor.

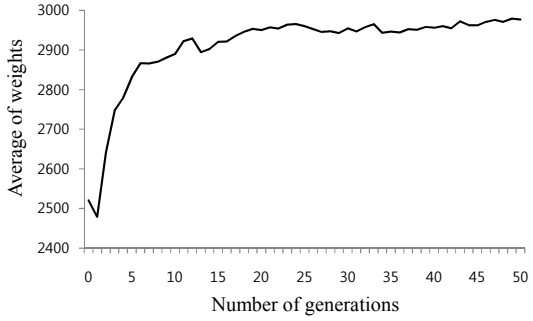


Fig 5. The change of hyperedge weights in the second layer's hypernetwork as generation goes on. Shown are the average weight rates of five hypernetworks.

B. The miRNA and mRNA Expression Classification

Figure 5 depicts the change of hyperedge weights in the second layer's hypernetworks as the generation goes on. The average weight curves are increased gradually, and stabilized after 50 epochs. The evolution process makes hyperedges having a specific pattern of data be survived for better classification. So, as the evolution progresses, the structure of hypernetworks is self-organized and finally converged to the optimal hypernetworks.

Table 2 presents the performance comparison of the proposed algorithm and other machine learning methods: support vector machines (SVMs), J48 decision trees (DTs), Bayesian networks (BNs), naïve Bayes (NB) and random forests (RFs) in WEKA. The linear kernel was used in SVMs with a sequential minimal optimization (SMO) and K2 is used as a search algorithm in Bayesian networks. All methods run 10 times using 10-fold cross validation, and averaged. As shown in Table 2, the LHN classifier shows 95.30% of accuracy. It is superior to decision trees, Bayesian networks, naïve Bayes and random forests, while providing competitive performance to the SVMs. Compared to the SVMs, the LHN classifiers yield interpretable results in addition to predictions.

TABLE 2. PERFORMANCE COMPARISON OF LAYERED HYPERNETWORKS AND CONVENTIONAL ALGORITHMS.

Algorithms	Avg. Accuracy (Stdev)
Support Vector Machines	97.30 (0.58)
Layered Hypernetworks	**95.39 (0.98)**
Bayesian Networks (# parents = 3)	92.92 (1.59)
Random Forests (# trees = 50)	90.90 (0.98)
Naïve Bayes	89.21 (0.79)
J48 Decision Trees	84.41 (2.70)

Accuracy denotes the ratio of the number of correctly classified samples to the number of total samples on a given test data set. Also, average are obtained after 10 times repeated experiments

TABLE 3. HIGH SCORED MIRNA-MRNA INTERACTIONS FORM LHNS.

No.	miRNAs	mRNAs (avg. scores)
I	*hsa-miR-21, has-miR-29a*	SFERS2 (959), NFIB (956), CRYAB (917), ALDH1A1 (916), CDKN1B (901), SART1 (864.5), HSPA1A (861.5), TGFBR2 (851.5), PPAP2B (837), PPP2CB (820), ARMCX1 (815.5), FLNB (814), HSP90B1 (797.5), GADD45G (795.5), MYLK (766), CTNNB1 (758.5), EPS8 (755), HNRPK (748.5), ZAK (745), FGFR1 (740.5), RRAGA (732.5), MT1X (730.5), CYP51A1 (728), BIRC4 (726), ZNF133 (713), GNA11 (712.5), YWHAE (712), ALDH2 (708.5), BNIP3L (708.5), TPM1 (704.5)
II	*hsa-miR-154, hsa-miR-184*	NFIB (824.5), SFRS2 (821.5), CRYAB (809), CDKN1B (792), ALDH1A1 (788), PPAP2B (772), HSPA1A (767), SART1 (756), PPP2CB (752), TGFBR2(737.5), FLNB (707), EPS8 (705.5)

C. Discovery of miRNA-mRNA Regulatory Modules

The proposed method is proper for the analysis of biological data, used to detect functional gene interactions. Table 3 shows the high scored miRNA-mRNA interactions found by the LHNs. Applying the proposed metric (significant score *S*), we extracted the significant relationships of miRNAs and mRNAs based on co-occurrence frequency in the combining of hyperedges of each modality. The significant scores are calculated from each miRNA and mRNA consisting of a module, and averaged. The miRNAs and mRNAs in Table 3 have been identified as cancer-related miRNAs and mRNAs in previous studies and their links have potential roles on the regulatory mechanism of diseases including cancers.

The hsa-miR-21 is up-regulated in various human cancers such as breast, cervical, colon, glioblastoma, hepatocellular, leukemia, lymphoma and ovarian carcinoma. Also, it is widely believed to have oncogenic activities by down-regulating the expression of many tumor suppressor genes such as PDCD4, PTEN and TPM1 [15], [16]. The over expression of hsa-miR-29a in lung cancer induces aberrant expression of methylation-silenced tumor suppressor genes and it has a functional role in human carcinoma cell invasion and proliferation [17], [18]. Also, hsa-miR-154 is expressed at significantly higher levels in human B cell chronic lymphocytic leukemia (CLL) cancer [19] and hsa-miR-184 has reported to be a putative suppressor of glioma progression [20]. In addition, most of the mRNAs in Table 3 are known to be associated with various human cancers and mRNAs of the same module share similar biological functions.

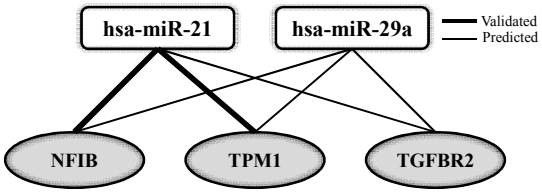
To verify the identified miRNA-mRNA interactions, we analyzed the functional coherences of mRNAs in the same module using Gene Ontology (GO). The GO has been a standard way to validate the functional correlations within a group of genes by a statistical significance analysis. If the identified miRNA-mRNA interactions imply a potential capacity, mRNAs consisting of modules may reflect their functional relevance in biological context.

We used the tool GOstat [21] to find significantly overrepresented GO terms. The GOstat calculates a *p*-value of annotated GO term based on the hypergeometric tests and performs the multiple comparison correction (*q*-value). We examined statistically significant GO terms with *q*-value < 0.01 for module I. The GO analysis results are shown in Table 4. Overall, the mRNAs in the module I belong to specific

TABLE 4. BIOLOGICAL PROCESSES OF POTENTIAL MODULE I ANNOTATED IN GO, OBTAINED BY GOSTAT (Q < 0.01, ADJUSTED *P*-VALUE).

GO ID	GO Terms	*q*-value
GO:0006915	Apoptosis	2.29E-03
GO:0012501	Programmed cell death	2.29E-03
GO:0008219	Cell death	2.29E-03
GO:0016265	Death	2.29E-03
GO:0042981	Regulation of apoptosis	3.44E-03
GO:0043057	Regulation of programmed cell death	3.44E-03
GO:0048869	Cellular developmental process	3.44E-03
GO:0030154	Cell differentiation	3.44E-03
GO:0048468	Cell development	3.45E-03
GO:0006916	Anti-apoptosis	4.07E-03

GO ID is the identification of the Gene Ontology (GO) term. *p*-value is calculated upon assuming hypergeometric distribution of annotated GO terms and multiple comparison correction (*q*-value)



miRNA	Chr.	Start-End Position	Strand
hsa-mir-21	Chr17	57918627-57918698	+
hsa-mir-29a	Chr 7	130561506-130561569	-

mRNA	Description
NFIB	Nuclear factor 1 B-type
TPM1	Tropomyosin alpha-1 chain
TGFBR2	Transforming growth factor, beta receptor II

Fig 6. Biologically significant a subset of miRNA-mRNA regulatory module with supporting evidence. The upper rectangles denote miRNAs and lower ovals are mRNAs. The shape of links corresponds to the biological relevance of each miRNA-mRNA binding event (Validated: experimentally supported target, Predicted: computationally predicted target). Tables represent the description of miRNAs and mRNAs comprising of a subset of module.

functional categories, which are related to cell mechanism such as apoptosis, cell death, cell differentiation, cell development and so on. These are all closely associated with cancer progression and development. This result suggests that the identified modules imply a specific functional role in cellular processes.

Fig. 6 shows the more biologically meaningful a subset of miRNA-mRNA regulatory module with supporting evidence is extracted from the miRNA and mRNA interactions in module I. Tables in Fig. 6 present the description of miRNAs and mRNAs consisting of module in detail. It shows the chromosomal location information and functional description of their shared putative mRNAs. As mentioned above, hsa-miR-21 and hsa-miR-29a affect cell growth and development, leading to a variety of disorders including human malignancies. And annotated genes, NFIB, TPM1 and TGFBR2, are closely associated with cancer. Especially, NFIB [22] and TPM1 [23] are experimentally tested target genes of has-miR-21 and actively involved in carcinogenesis. The nuclear factor I/BI (NFIB) binds the hsa-miR-21 promoter as a negative regulator and the repression of NFIB by hsa-miR-21 causes a translational suppression in the biological process. Tropomyosin 1 (TPM1) as a tumor suppressor is also implicated in cell migration and invasion and the down-regulation of TPM1 by hsa-miR-21 leads to tumor growth and progression. In addition, transforming growth factor beta 2 (TGFBR2) is an essential regulator of cellular processes including proliferation, migration and cell survival and TGFBR-mediated inhibition of proliferation is frequently observed in human cancer [24]. As a result, we can conclude the proposed method discover biologically valuable miRNA-mRNA modules with strong correlation which evidently relate to cancer mechanism.

VI. Conclusions

We propose a novel method, layered hypernetworks (LHNs), and apply the model to identifying functional miRNA-mRNA regulatory modules effectively from expression profiles. An evolutionary strategy is drawn to detect the best combinations of higher-order building blocks without exhaustive search in limited computing sources. The LHN structure is appropriate for detecting biologically relevant gene interactions from heterogeneous information sources, because it can represent the relationships between more than one modality. Also, the LHN can provide interpretable results for understanding of synergistic relations among multiple resources and can produces good classification performance.

In this study, the proposed method is applied to the miRNA and mRNA expression profiles data associated with multiple human cancers. The experimental results show our method outperforms decision trees, Bayesian networks, naïve Bayes and random forests, while having competitive performance to support vector machines. The results also show that the LHN can identify miRNA-mRNA interactions with biological relevance. The discovered miRNA-mRNA regulatory modules are validated by GO analysis and literature reviews.

In this study, the procedure for generating and learning a LHN is based on the random selection. Future work includes the introduction of prior knowledge such as correlation information in the LHN frameworks to obtain more biologically meaningful results.

Acknowledgment

S.-J. Kim thanks In-Hee Lee and Jae-Hong Eom for discussion.

References

[1] J.-W. Nam, J. Kim, S.-K. Kim, and B.-T. Zhang, ProMiR II: a web server for the probabilistic prediction of clustered, nonclustered, conserved and nonconserved microRNAs, *Nucleic Acid Research*, 34, pp. W455- W458, 2006.

[2] B.P. Lewis, C.B. Burge, and D.P. Bartel, Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, *Cell*, 120(1), pp. 15-20, 2005.

[3] S. Yoon and G. De Micheli, Prediction of regulatory modules comprising microRNAs and target genes, *Bioinformatics*, 21, pp. ii93-ii100, 2005.

[4] J.C. Huang, Q.D. Morris, and B.J. Frey, Detecting microRNA targets by linking sequence microRNA and gene expression data*, In Proceedings of RECOMB*, pp. 114-129, 2006.

[5] J.-G Joung, K.-B. Hwang, J.-W. Nam, S.-J. Kim, and B.-T. Zhang, Discovery of microRNA-mRNA modules via population-based probabilistic learning, *Bioinformatics*, 23(9), pp. 1141-1147, 2007.

[6] B.-T. Zhang, Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory, *IEEE Computational Intelligence Magazine*, **3**(3). pp. 49-63, 2008.

[7] J. Lu, G. Getz, E.A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B.L. Ebert, R. H. Mak, A.A. Ferrando, J.R. Downing, T. Jacks, H.R. Horvitz, and T.R. Golub, MicroRNA expression profiles classify human cancers, *Nature*, 435(7043), pp. 834-848, 2005.

[8] S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C.-H. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, J. P. Mesirov, T. Poggio, W. Gerald, M. Loda, E. S. Lander, and T. R. Golub, Multiclass cancer diagnosis using tumor gene expression signature, *PNAS*, 98(26), pp. 15149-15154, 2001.

[9] J.-K. Kim and B.-T. Zhang, Evolving hypernetworks for pattern classification, *IEEE Congress on Evolutionary Computation (CEC 2007)*, pp. 1856-1862, 2007.

[10] J.-W. Ha, J.-H. Eom, S.-C. Kim, and B.-T. Zhang, Evolutionary hypernetwork models for aptamer-based cardiovascular disease diagnosis, *The Genetic and Evolutionary Computation Conference (GECCO 2007)*, pp. 2709-2716, 2007.

[11] S. Kim, S.-J. Kim, and B.-T. Zhang, Evolving hypernetwork classifiers for microRNA expression profile analysis, *IEEE Congress on Evolutionary Computation (CEC 2007)*, pp. 313-319, 2007.

[12] E. Bautu, S. Kim, A. Bautu, H. Luchian, and B.-T. Zhang, Evolving hypernetwork models of binary time series for forecasting price movements on stock markets, *IEEE Congress on Evolutionary Computation (CEC 2009)*, pp. 166-173, 2009.

[13] P.A. Futreal, L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, and M.R. Stratton, A census of human cancer genes, *Nature Reviews Cancer*, 4, pp. 177–183, 2004.

[14] Q. Cui, Y. Ma, M. Jaramillo, H. Bari, A. Awan, S. Yang, S. Zhang, L. Liu, M. Lu, M. O'Connor-McCourt, E.O. Purisima, and E. Wang, A map of human cancer signaling, *Molecular Systems Biolology*, 3(152), 2007.

[15] S. Zhu, H. Wu, F. Wu, D. Nie1, S. Sheng and Y.-Y. Mo, MicroRNA-21 targets tumor suppressor genes in invasion and metastasis, *Cell Research*, 18, pp. 350-359, 2008.

[16] M.L. Si, S. Zhu, H. Wu, Z. Lu, F. Wu, and Y.Y. Mo, miR-21-mediated tumor growth, *Oncogene*, 26(19), pp. 2799-2803, 2007.

[17] M. Fabbri, R. Garzon, A. Cimmino, Z. Liu, N. Zanesi, E. Callegari, S. Liu, H. Alder, S. Costinean, C. Fernandez-Cymering, S. Volinia, G. Guler, C.D. Morrison, K.K. Chan, G. Marcucci, G.A. Calin, K. Huebner, and C.M. Croce, MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B, *PNAS*, 104(40), pp. 15805-15810, 2007.

[18] M.K. Muniyappa, P. Dowling, M. Henry, P. Meleady, P. Doolan, P. Gammell, M. Clynes, and N. Barron, MiRNA-29a regulates the expression of numerous proteins and reduces the invasiveness and proliferation of human carcinoma cell lines, *Eur J Cancer*, 45(170), pp. 3104-3118, 2009.

[19] G.A. Calin, C.G. Liu, C. Sevignani, M. Ferracin, N. Felli, C.D. Dumitru, M. Shimizu, A. Cimmino, S. Zupo, M. Dono, M.L. Dell'Aquila, H. Alder, L. Rassenti, T.J. Kipps, F. Bullrich, M. Negrini, and C.M. Croce, MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias, *PNAS*, 101(32), pp. 11755-11760, 2004.

[20] B. Malzkorn, M. Wolter, F. Liesenberg, M. Grzendowski, K. Stühler, H.E. Meyer, and G. Reifenberger, Identification and functional characterization of microRNAs involved in the malignant progression of gliomas, *Brain Pathology*, 2009.

[21] T. Beissbarth and T.P. Speed, GOstat: Find statistically overrepresented Gene Ontologies within a group of genes, *Bioinformatics*, 20(9), pp. 1464-1465, 2004.

[22] S. Fujita, T. Ito, T. Mizutani, S. Minoguchi, N. Yamamichi, K. Sakurai, and H. Iba, miR-21 gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism, *Journal of Molecular Biology*, 378(3), pp. 492-504, 2008.

[23] S. Zhu,M.L. Si, H. Wu, and Y.Y. Mo, MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1), *Journal of Biological Chemistry*, 282(19), pp. 14328-14336, 2007.

[24] R.L. Elliott and G.C. Blobe, Role of transforming growth factor beta in human cancer, *Journal of Clinical Oncology*, 23(9), pp. 2078-2093, 2005.