# Evolutionarily Learning Multi-aspect Interactions and Influences from Network Structure and Node Content

**Songlei Jian[1,2], Liang Hu[2,3], Longbing Cao[2], Kai Lu[1], Hang Gao[4]**
[1]College of Computer, National University of Defense Technology, China
[2]Advanced Analytics Institute, University of Technology Sydney, Australia
[3]University of Shanghai for Science and Technology, [4]Alibaba Group, China
jiansonglei@163.com, rainmilk@gmail.com, longbing.cao@uts.edu.au,
kailu@nudt.edu.cn, allen.gh@alibaba-inc.com

## Abstract

The formation of a complex network is highly driven by multi-aspect node influences and interactions, reflected on network structures and the content embodied in network nodes. Limited work has jointly modeled all these aspects, which typically focuses on topological structures but overlooks the heterogeneous interactions behind node linkage and contributions of node content to the interactive heterogeneities. Here, we propose a multi-aspect interaction and influence-unified evolutionary coupled system (MAI-ECS) for network representation by involving node content and linkage-based network structure. MAI-ECS jointly and iteratively learns two systems: a multi-aspect interaction learning system to capture heterogeneous hidden interactions between nodes and an influence propagation system to capture multi-aspect node influences and their propagation between nodes. MAI-ECS couples, unifies and optimizes the two systems toward an effective representation of explicit node content and network structure, and implicit node interactions and influences. MAI-ECS shows superior performance in node classification and link prediction in comparison with the state-of-the-art methods on two real-world datasets. Further, we demonstrate the semantic interpretability of the results generated by MAI-ECS.

## Introduction

Complex networks are ubiquitous; precisely representing a complex network is critical for understanding and developing networking applications yet challenging due to intricate network structures, node information and influence, and interactions between nodes (Cui et al. 2018; Cai, Zheng, and Chang 2017). As the citation network in Figure 1 illustrates, each paper $p$ owns its content information, including title, abstract and other information, and plays various influences on other papers w.r.t. such aspects as research topic uniqueness, method novelty, and experimental advantage. Accordingly, different citation network structures are formed under different scenarios, as shown in Figure 1(a) and Figure 1(b). While the network structures reflect different node connections (citations) between three context papers $p_1$, $p_2$ and $p_3$ and the target paper $p_t$ in scenarios (a) and (b), it is important to understand the underlying node interactions and the roles of node information in influencing node interactions, which are hidden behind each network structure. For example, the influences of each paper and the multi-aspect interactions between context papers and between the target paper and context papers in scenario (a) may be quite different from that in scenario (b), even though both scenarios share the same node content of each paper. It is important for us to understand the multi-aspect node influences and the multi-aspect interactions between nodes that drive the formation of connections.

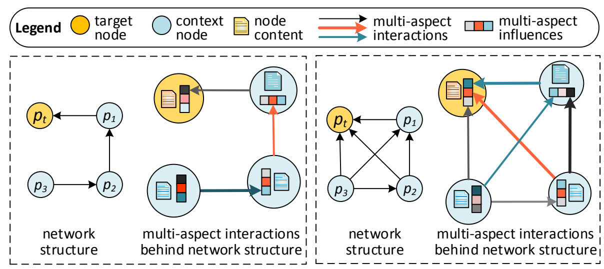



Figure 1: Two scenarios in citation networks: different network structures driving by different node influences and interactions between nodes.

The above example illustrates a critical perspective in complex networks and systems, i.e., representing multi-aspect and heterogeneous interactions and influences between nodes (objects), in order to understand some intrinsic characteristics and fundamental complexities: explicit and implicit heterogeneities and coupling relations in complex networks and systems (Cao 2014; Cao 2015; Zhang et al. 2018a). This requires to involve, model and integrate (1) explicit and heterogeneous sources of node content information (e.g., paper's title and/or abstract) and network topological structure (e.g., paper citations), and (2) implicit and heterogeneous aspects of node influences (e.g., a paper's topic uniqueness and design novelty) and node interactions (e.g., a paper cites the algorithm introduced in another paper). However, the above significantly challenges existing research on network embedding and representation. To the best of our knowledge, no work has been reported to jointly disclose and model these aspects.

Most of existing plain network representation methods (Li et al. 2014; Tang et al. 2015; Grover and Leskovec 2016) aim to preserve the network structure. Some recent methods tend to integrate topological and content information by extending existing plain network representation methods. For example, TriDNR (Pan et al. 2016) conducts random walk on node text to build an independent term in its objective function. However, it models the network structure and node content separately through a weighted combination, which fails to jointly consider the information from network structure and node content. Although CANE (Tu et al. 2017) builds context-aware network embedding, it only models the different semantic aspects of node content without considering the multi-aspect node influence or the multi-aspect interactions between nodes. The work in (Gao and Huang 2018) uses two Autoencoders on network structure and node content separately and combines them through alignment on their consistent information. However, the above methods overlook the multi-aspect interactions between node pairs or only model the different attentions within node content.

To model the explicit and implicit information and multi-aspect heterogeneous interactions discussed above in a complex network, we here propose a multi-aspect interaction and influence-unified evolutionary coupled system (MAI-ECS), as shown in Figure 2. MAI-ECS involves both node content and network structure information, and couples a multi-aspect interaction learning system (MAI-LS) to model the hidden and various interactions between nodes with a multi-aspect influence propagation system (MAI-PS) to generate the implicit and multi-aspect node influences. The MAI-LS decomposes a node linkage (an edge in a network) into several hidden interactive relations to explain why the linkage forms, and then MAI-PS discloses a node's various influences on others in terms of node attributes (e.g., content) and its multiple roles in forming a topological structure. The two coupled systems (Geiser 2014) in MAI-ECS are alternately optimized to form an evolutionary learning system, until producing the desired network representation.

We apply MAI-ECS to node classification and link prediction on two real-world data sets, showing significant performance improvement over the start-of-the-art baselines. We further demonstrate the semantic meaning of learned multi-aspect influence and multi-aspect interactions, indicating the interpretability of MAI-ECS.

## Related Work

**Plain Network Representation** A typical solution to represent nodes or edges is to factorize graph Laplacian eigenmaps or the node proximity matrix directly (Yin, Gao, and Lin 2016; He and Niyogi 2004). With deep learning, recently a lot of methods try to incorporate deep learning models to learn node representation. Among them, the most representative way is to combine the random walk and word embedding, such as Skip-gram and CBOW (Mikolov et al. 2013). DeepWalk (Perozzi, Al-Rfou, and Skiena 2014) and Node2vec (Grover and Leskovec 2016) build context through the random walk and feed the contexts into a skip-gram model. LINE (Tang et al. 2015) and SDNE (Wang, Cui, and Zhu 2016) tend to learn representations from local network structure while other methods aim to capture the global structure and community patterns, such as GraRep (Cao, Lu, and Xu 2015) and M-NMF (Wang et al. 2017). Other method (Hu et al. 2019) considers the context influence in multiple networks and applies into recommender systems. Only considering the structure information cannot capture the implicit and heterogeneous aspects of node influence and interactions. Further, the random walk and deep learning model are treated as two independent parts in these methods. This means the two parts cannot optimize each other, which is not suitable for jointly learning node content and network structure.

**Attributed Network Representation** Due to the heterogeneity and couplings (Cao 2014) between node content information and network topological information, only a limited number of methods are able to import content information into network representation. Text-associated DeepWalk (TADW) (Yang et al. 2015) firstly attempts to incorporate node text information into node representation through matrix factorization. TriNDR (Pan et al. 2016)] further exploits node labels to make node representations better for node classification. The method in (Zhang et al. 2018b) also models the node text through attribute-aware Skip-gram model. Then the node text and node structure information are combined w.r.t. a weight parameter, which is largely affected by the performance and generalizability of the method. UPP-SNE (Zhang et al. 2017) incorporates user profile with network structure by mapping them into the same space. In this work, the user profile attributes and the network structure are assumed highly interrelated so the nodes in a context are represented by their attribute vectors. All these methods also conduct random walk to collect contexts and use a deep model to learn representation. The node content information is forced to fit the topological structure without considering the multi-aspect node influence or interactions brought by node attributes. CANE (Tu et al. 2017) considers the node context's influence on the node content information and use mutual attention mechanism to build context-aware embedding. However, only modeling the aspects in content cannot capture the implicit aspects of node influence or interactions. The method in (Gao and Huang 2018) tries to modify autoencoder to align the structure information and node attributes, which only keeps the consistent information and leads to information loss.

## Problem Formalization

A network with node content is denoted as $G = (\mathcal{V}, \mathcal{E}, \mathcal{C})$, where $\mathcal{V} = \{v_i\}_{i=1}^{N}$ denotes the node set, $e_{i,j} \in \mathcal{E}$ denotes an edge, $(v_i, v_j, w_{ij})$, with weight $w_{ij}$ between source node $v_i$ and target node $v_j$, and $\mathbf{c}_i \in \mathcal{C}$ is $v_i$'s content information which may be a text, an image or an attribute vector. The MAI-ECS $\mathcal{M}_\Theta$ aims to learn a network representation $\mathcal{N} = \{\mathbf{Y}, \bar{\mathbf{S}}, \mathbf{R}; \mathbf{\Theta}\}$ where $\mathbf{y}_i \in \mathbf{Y}$ is a low-dimensional vector (with the length $K$) for each node, $\bar{\mathbf{s}}_i \in \bar{\mathbf{S}}$ is the stable multi-aspect influence state vector with $L$ aspects ($\bar{\mathbf{s}}_i \in [0, 1]^L$)

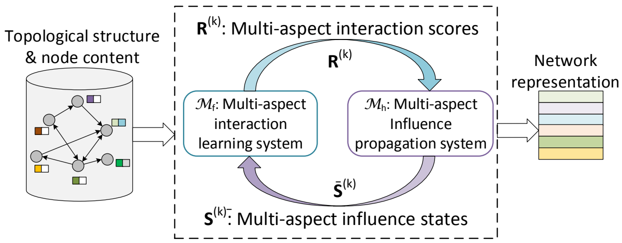



Figure 2: The MAI-ECS for jointly learning multi-aspect influences and interactions in a complex network.

for each node, $\mathbf{r}^{i\to j} \in \mathbf{R}$ denotes the multi-aspect influence scores for each edge, and $\mathbf{\Theta}$ denotes model parameters.

$\mathcal{M}_\Theta$ consists of two coupled systems: $\mathcal{M}_h$ is a multi-aspect influence propagation system based on the transition configuration $\mathbf{R}$ fed by $\mathcal{M}_f$; $\mathcal{M}_f$ is a multi-aspect interaction learning system to refine the configuration $\mathbf{R}$ according to the stable influential states $\bar{\mathbf{S}}$ fed by $\mathcal{M}_h$. We aim to obtain the desired network representation $\mathcal{N}$ via an evolutionary learning process over $\mathcal{M}_h$ and $\mathcal{M}_f$.

## Multi-aspect Evolutionary Coupled Systems

Here, we present the details of the coupled evolution process to learn the underlying multi-aspect interactions. As illustrated in Figure 2, the learning process of the coupled systems can be formalized as the following evolutionary equations, where $k$ denotes the evolutionary iteration:

$$\begin{cases} \mathcal{M}_f: & \mathbf{R}^{(k+1)} = f_{\mathbf{\Theta}}(\bar{\mathbf{S}}^{(k)}, \mathbf{R}^{(k)}|G) \quad (1) \\ \mathcal{M}_h: & \bar{\mathbf{S}}^{(k+1)} = h(\mathbf{R}^{(k+1)}, \bar{\mathbf{S}}^{(k)}|G) \quad (2) \end{cases}$$

The above two coupled equations outline the evolutionary dynamics between $\mathcal{M}_f$ and $\mathcal{M}_h$. MAI-LS $\mathcal{M}_f$ aims to learn multi-aspect interaction scores $\mathbf{R}^{(k)}$ for all edges, which is largely dependent on the multi-aspect influence states $\bar{\mathbf{S}}^{(k)}$ for all nodes. In turn, MAI-PS $\mathcal{M}_h$ updates the stable multi-aspect influence states $\bar{\mathbf{S}}^{(k+1)}$ according to the refined transition configuration based on $\mathbf{R}^{(k)}$. Therefore, this is an asymptotic evolution process over the coupled systems $\mathcal{M}_f$ and $\mathcal{M}_h$ to learn the optimized network representation $\mathcal{N}$. In the following subsections, we introduce $\mathcal{M}_f$ and $\mathcal{M}_h$, and the evolutionary learning algorithm based on these two coupled systems.

### Multi-aspect Interaction Learning System

According to Eq. 1, MAI-LS takes the latest stable multi-aspect influence states $\bar{\mathbf{S}}^{(k)}$ as the input to learn the multi-aspect interaction scores, which consists of two submodels: node representation model and interaction scoring model.

**Node Representation Model** Without loss of generality, we present the node content w.r.t. textual data. Other content can be represented by their representations, e.g., VGG features (Simonyan and Zisserman 2014) for images. First, we map each word $j$ in a node text content $\mathbf{c}_i$ into a pre-trained word embedding $\mathbf{w}_j \in \mathbb{R}^{K\times 1}$ (Pennington, Socher, and Manning 2014). Then, the text embedding $\mathbf{t}_i$ is encoded by the attention mechanism (Hu et al. 2018):

$$\mathbf{t}_i = \sum_{\mathbf{w}_j \in \mathbf{c}_i} \alpha_j \mathbf{w}_j \quad (3)$$

where the attention weight $\alpha_j$ is learned through a two-layer attention network:

$$\alpha_j = \mathrm{softmax}(\tanh(\mathbf{W}^t \mathbf{t} + b^t)) \quad (4)$$

$\mathrm{softmax}(x_k) = e^{x_k} / \sum_j e^{x_j}$ and $\mathbf{W}^t \in \mathbb{R}^{1\times K}$ and $b^t \in \mathbb{R}$ is the bias term.

Then, we associate an embedding vector $\mathbf{e}_i \in \mathbb{R}^{K\times 1}$ for each node $v_i$ to learn the topological structure information. Since the content embedding $\mathbf{t}_i$ and structure embedding $\mathbf{e}_i$ are characterized by different properties, we combine them as the node representation $\mathbf{y}_i$ to capture more comprehensive information in the network:

$$\mathbf{y}_i = L_2 norm(\mathbf{e}_i \oplus \mathbf{t}_i) \quad (5)$$

where $L_2 norm$ denotes $L_2$ normalization and $\oplus$ denotes a merging operation, e.g., concatenation. In this paper, we adopt the addition operation which results in the best performance in our experiments.

**Interaction Scoring Model** As discussed in the introduction, each observed edge $e_{i,j}$ can be regarded as the formation of the underlying interactions between a node pair $(v_i, v_j)$ w.r.t. different aspects. The multi-aspect interaction scores are relevant to the properties of interactive node pair $(v_i, v_j)$, which are encoded by the node representations $(\mathbf{y}_i, \mathbf{y}_j)$, and their influential states $(\bar{\mathbf{s}}_i, \bar{\mathbf{s}}_j)$, which are generated by MAI-PS to depict the interaction aspects and their strengths.

First, we measure the compatibility of interactions according to the node properties encoded by $\mathbf{y}_i$ and $\mathbf{y}_j$:

$$\mathbf{y}_{i,j} = \mathbf{y}_i \odot \mathbf{y}_j \quad (6)$$

where $\odot$ denotes the element-wise product.

Then, we encode the attraction of node $v_j$ on node $v_i$ for directed interaction and mutual attraction between $v_j$ and $v_i$ for undirected interaction according to the nodes' influential states $\bar{\mathbf{s}}_i$ and $\bar{\mathbf{s}}_j$:

$$\mathbf{x}_{i,j} = \begin{cases} \mathbf{W}^s \bar{\mathbf{s}}_j & \text{if } (v_i, v_j) \text{ is directed} \\ \mathbf{W}^s \bar{\mathbf{s}}_i \odot \mathbf{W}^s \bar{\mathbf{s}}_j & \text{otherwise} \end{cases} \quad (7)$$

where $\mathbf{W}^s \in \mathbb{R}^{L\times L}$. In the case of directed interactions, the formation of an interaction is mainly relevant to the influence of the target node, e.g., a citation is formed due to the influence of its target paper. In the case of an undirected or bi-directional interaction, its formation is driven by the influence of both nodes, e.g., the collaboration between two influential authors.

As a result, the overall interaction score $\mathcal{S}^O_{i,j}$ between $(v_i, v_j)$ is measured by summing over $\mathbf{y}_{i,j}$ and $\mathbf{x}_{i,j}$:

$$\mathcal{S}^O_{i,j} = \sum \mathbf{y}_{i,j} + \sum \mathbf{x}_{i,j} = \mathbf{y}_i^\top \mathbf{y}_j + \bar{\mathbf{s}}_i^\top \mathbf{W}^{s\top} \mathbf{W}^s \bar{\mathbf{s}}_j \quad (8)$$

Moreover, we need to determine the interaction aspect between node pair $(v_i, v_j)$. First, we score the interaction for each aspect as follows:

$$\mathcal{S}^A_{l,i,j} = \mathbf{W}^{y\top}_l \mathbf{y}_{i,j} + \mathbf{W}^{x\top}_l \mathbf{x}_{i,j} + b_l \tag{9}$$

where $\mathbf{W}^y_l \in \mathbb{R}^{K\times 1}$, $\mathbf{W}^x_l \in \mathbb{R}^{L\times 1}$ are the parameters for measuring the score of aspect $l$. In the same way, we obtain the scores for all aspects.

These aspects of scores imply the probability of the interaction aspects, which can be viewed as a mixture model of the interactions over multiple aspects. As a result, the assignment of the interaction aspects, $a_{i,j} \in \{1, \cdots, L\}$, can be drawn from the following categorical distribution:

$$a_{i,j} \sim \boldsymbol{\pi}(\{\mathcal{S}^A_{l,i,j}\}_{l=1,\cdots,L}) \tag{10}$$

where the probability mass function is defined as:

$$\pi_l = \text{softmax}(\tanh(\mathcal{S}^A_{l,i,j})), \quad l = 1, \cdots, L \tag{11}$$

However, the assignment $a_{i,j}$ is a discrete random variable which cannot work with other continuous variables due to the absence of the gradient for backpropagation (Jang, Gu, and Poole 2016) in a neural network. As a workaround, we adopt the Gumbel-SoftMax (Jang, Gu, and Poole 2016) to conduct a soft sampling for discrete values:

$$a_{l,i,j} = \frac{\exp\left((\log \pi_l + g_l)/\tau\right)}{\sum_{k=1}^{L} \exp\left((\log \pi_k + g_k)/\tau\right)} \tag{12}$$

where $g = -\log(-\log(u))$ and $u \sim uniform(0, 1)$. In this paper, we set the temperature parameter $\tau = 0.1$, which tends to output a one-hot-like assignment, that is, only one aspect has the mass close to 1 and the masses of other aspects are close to 0. Accordingly, we get a relaxed discrete version of multi-aspect interaction scores $\mathbf{r}^{i\to j} \in \mathbf{R}$, where only one aspect score tends to be preserved:

$$r^{i\to j}_l = \max(a_{l,i,j}\mathcal{S}^A_{l,i,j}, 0), \quad l = 1, \cdots, L \tag{13}$$

## Multi-aspect Influence Propagation System

So far we have learned multi-aspect interaction scores $\mathbf{R}^{(k)}$ (cf. Eq. 13) from MAI-LS $\mathcal{M}_f$. Correspondingly, the multi-aspect influence of each node is changed with the updated multi-aspect interactions in the network. According to Eq. 2, MAI-PS $\mathcal{M}_h$ takes $\mathbf{R}^{(k)}$ to configure the influence propagation dynamical system. This dynamical system is defined in terms of the following state space equations:

$$\mathbf{S}^{(k)}_{:,l}(n+1) = \mathbf{T}_l \mathbf{S}^{(k)}_{:,l}(n) \tag{14}$$

$$\mathbf{S}^{(k)}_{:,l}(0) = \bar{\mathbf{S}}^{(k-1)}_{:,l}. \tag{15}$$

$\mathbf{S}^{(k)}_{:,l}(n)$ denotes the influential states w.r.t. aspect $l$ for all nodes, where $n$ denotes the propagation time step in the dynamical system. Since $\mathcal{M}_h$ is a multi-aspect influence propagation system, we have $L$ transition matrices $\{\mathbf{T}_l\}^L_{l=1}$.

Given an aspect $l$, transition matrix $\mathbf{T}_l \in [0, 1]^{N\times N}$ is configured as follows:

$$\mathbf{T}_l = \frac{(1-d)}{N}\mathbf{1}\mathbf{1}^\top + d(\mathbf{P}_l + \mathbf{D}_l) \tag{16}$$

where $\mathbf{1} \in \mathbb{R}^{N\times 1}$ is all-one vector and $d$ is the damping ratio to prevent sticking in local network neighborhoods. We set $d = 0.95$ in the experiments. The matrix $\mathbf{D}_l$ is set to handle the dangling nodes that do not have any outlink in aspect $l$:

$$\mathbf{D}_{l[:,i]} = \begin{cases} \frac{1}{N}\mathbf{1}, & \text{if } v_i \text{ is dangling node} \\ \mathbf{0}, & \text{otherwise.} \end{cases} \tag{17}$$

$\mathbf{P}_l$ is configured according to multi-aspect interaction scores $\mathbf{R}^{(k)}$. The transition probability $p^{i\to j}$ from $v_i$ to $v_j$ is proportional to the corresponding $r^{i\to j}_l$.

$$p^{i\to j} = \begin{cases} \frac{r^{i\to j}_l}{\sum_{i=1}^{N} r^{i\to j}_l}, & \text{if } (v_i, v_j) \in \mathcal{E} \text{ and } a_{i,j} = l \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

where $a_{i,j} = l$ means that we only consider the interaction on the sampled interaction aspect (cf. Eq. 10) since $r^{i\to j}_l$ is a softly sampled score (cf. Eq. 13). As a result, we have $\mathbf{P}_l$:

$$\mathbf{P}_l = \begin{bmatrix} p^{1\to 1} & p^{2\to 1} & \dots & p^{N\to 1} \\ p^{1\to 2} & \ddots & & \vdots \\ \vdots & & p^{i\to j} & \\ p^{1\to N} & \dots & & p^{N\to N} \end{bmatrix}. \tag{19}$$

Under the above settings for $\mathbf{T}_l$, this influence propagation dynamical system sets up a Markov chain for each aspect $l$. According to Markov's theorem: *a Markov chain is ergodic if there is a positive probability to pass from any state to any other state in one step.* Obviously, $\mathbf{T}_l$ satisfies the condition of ergodicity, which guarantees the stable state $\bar{\mathbf{S}}_{:,l}$ for each influential aspect within finite time steps.

Eq. 14 implies the power method (Saad 2011) to find $\bar{\mathbf{S}}$. Since $\mathbf{P}_l$ is very sparse, the total number of nonzero entries over all $\{\mathbf{P}_l\}$ is $|\mathcal{E}|$, i.e., the number of edges, the influence state matrix $\mathbf{S}(n)$ for all aspects can be computed in $O(|\mathcal{E}|)$ CPU time or in $O(1)$ GPU time for each power iteration.

## Evolutionary Learning Process

The coupled systems MAI-LS $\mathcal{M}_f$ and MAI-PS $\mathcal{M}_h$ set up the evolutionary dynamics over the network. Notice that an observed structure of a network is evolved from some previous structure with fewer edges. That is, given a masked network structure with masked edges $\mathcal{E}^- \subset \mathcal{E}$, the evolutionary dynamics tend to drive it to recover the target network structure, as demonstrated in Figure 3. Inspired by this idea, we set up an evolutionary objective for MAI-LS $\mathcal{M}_f$ and MAI-PS $\mathcal{M}_h$ based on sampled masked network structures.

Given a masked network structure, the evolutionary direction is driven by the interaction between nodes. Therefore, given a source node $v_i$ and any two other nodes $v_j, v_k$, we have the following orders of the interaction strengths:

$$\begin{cases} \mathcal{O}_{i,j} \succ \mathcal{O}_{i,k}, & \text{if } w_{i,j} > w_{i,k} & (20) \\ \mathcal{O}_{i,j} \succ \mathcal{O}_{i,k}, & \text{if } (v_i, v_j) \in \mathcal{E} \text{ and } (v_i, v_k) \notin \mathcal{E} & (21) \\ \mathcal{O}_{i,j} \succ \mathcal{O}_{i,k}, & \text{if } (v_i, v_j) \in \mathcal{E}^- \text{ and } (v_i, v_k) \notin \mathcal{E} & (22) \end{cases}$$

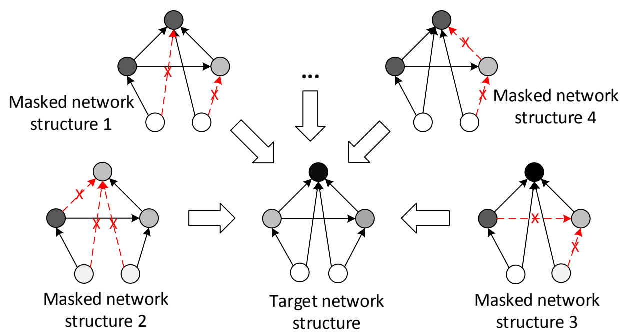



Figure 3: The evolutionary learning process.

Eq. 20 specifies the interaction order according to the weight of edges; Eq. 21 specifies the interaction order according to the presence of edges; and Eq. 22 specifies the interaction order according to the possibility to form new edges.

According to the above order relationships, we can set up the following objective over $\langle v_i, v_j, v_k \rangle$:

$$\begin{cases} S^O_{i,j} > S^O_{i,k}, & \text{if } \mathcal{O}_{i,j} \succ \mathcal{O}_{i,k} \quad (23) \\ S^A_{i,j}|a_{l,i,j} > S^A_{i,k}|a_{l,i,j}, & \text{if } \mathcal{O}_{i,j} \succ \mathcal{O}_{i,k} \quad (24) \end{cases}$$

where $S^O_{i,j}$ is overall interaction score given by Eq. 8 and $S^A_{i,j}|a_{l,i,j} = \sum a_{l,i,j} S^A_{l,i,j}$, $S^A_{i,k}|a_{l,i,j} = \sum a_{l,i,j} S^A_{l,i,k}$ are the aspect-specific interaction scores conditional on the sampled interaction aspect (cf. Eq. 12). Accordingly, we set up the following triplet max-margin losses:

$$\begin{cases} L_O = \max\{0, \varepsilon_O - (S^O_{i,j} - S^O_{i,k})\} \quad (25) \\ L_A = \max\{0, \varepsilon_A - (S^A_{i,j}|a_{l,i,j} - S^A_{i,k}|a_{l,i,j})\} \quad (26) \end{cases}$$

where $\varepsilon_O$ and $\varepsilon_A$ are the max-margin parameters, we set both of them to 1 in the experiments. Specifically, $L_O$ mainly aims to optimize the overall topological interaction resulting in the edge formation while $L_A$ mainly aims to optimize the semantic interaction in a specific aspect resulting in the edge formation.

We use mini-batch learning to learn the parameters $\mathbf{\Theta}$ of $\mathcal{M}_f$. Then the mean loss of a mini-batch $\mathcal{B}$ for computing gradients is given as:

$$\mathcal{L} = \frac{1}{|\mathcal{B}|} \sum_{\langle v_i, v_j, v_k \rangle \in \mathcal{B}} (L_O + L_A) \tag{27}$$

We adopt Adam (Ruder 2016) for optimizing the gradients $\nabla_{\mathbf{\Theta}} \mathcal{L}$ to find the optimal model parameters $\mathbf{\Theta}$.

The whole learning process is depicted in Algorithm 1. For the overall iteration, we randomly sample a masked network structure to learn the coupled systems $\mathcal{M}_f$ and $\mathcal{M}_h$ in turn. This process begins with $\mathcal{M}_f$ with randomly initialized $\bar{\mathbf{S}}^{(0)}$. In $\mathcal{M}_h$, since the dynamical system is ergodic, we employ the power method on GPU to find the stable influential states. The MAI-ECS code is available at: https://github.com/jiansonglei/MAI-ECS.

**Algorithm 1** The evolutionary learning process of MAI-ECS

**Require:** $G$: network, $M_1$: maximum iterations of evolutionary learning process, $M_2$: maximum iterations in $\mathcal{M}_f$, $M_3$: maximum iterations in $\mathcal{M}_h$

**Ensure:** $\mathcal{N} = \{\mathbf{Y}, \bar{\mathbf{S}}, \mathbf{R}; \mathbf{\Theta}\}$ - the network representation

```
Initialize S̄^(0) with a random value
for k = 0, ··· , M_1 do
    Sample a masked network G_k from G
    R^(k+1) = f_Θ(S̄^(k), R^(k) | G_k)
    S̄^(k+1) = h(R^(k+1), S̄^(k) | G_k)
end for
Construct Y according to Eq. 5 and extract Θ from M_f
return Y, S̄, R, Θ
function f_Θ(S̄, R | G)
    while iter ≤ M_2 do
        Optimize ∇_Θ L (cf. Eq.27) with Adam
    end while
    Generate r^(i→j) according to Eq. 13
    return R, Θ
end function
function h(R, S̄ | G )
    Construct {T_l}_{l=1}^L according to Eq. 16
    while n ≤ M_3 do                          ▷ Power iteration
        S_{:,l}(n + 1) = T_l S_{:,l}(n) for l ∈ [1, ··· L]
    end while
    return S̄
end function
```

## Experiments

In this section, we evaluate our proposed model on two real-world datasets and two applications, and compare it with the state-of-the-art network representation methods.

### Experimental Setup

The citation network Aminer[1] collects data from DBLP, ACM, MAG (Microsoft Academic Graph), and other sources. The nodes denote the papers and the edge denotes the citation relationship which is directed. The whole network contains millions of papers from various domains. We choose papers from four domains, i.e., artificial intelligence, database & data mining, theoretical computer science, and computer graphics & multimedia. The domains are used for labels in node classification. For each paper, we use its title as the node content information. The final citation network contains 29,896 nodes and 66,166 edges.

The second dataset is a coauthoring network in Aminer [2], where the nodes denote authors, the edges denotes the coauthorships and the edge weight denotes the number of coauthored papers. For each author, we use her/his research interests as the node content information. This network contains 1,712,433 authors and 4,258,615 edges.

In our experiments, we evaluate the network representation on standard learning tasks: node classification on cita-

[1] https://aminer.org/citation (V7 version is used)

[2] https://aminer.org/aminernetwork

Table 1: Node classification performance on citation network

| $p\%$ | Metric | Node2vec | Doc2vec | NV+DV | TriDNR | CANE | MAI-LS | MAI-ECS |
|---|---|---|---|---|---|---|---|---|
| 80 | Accuracy | 0.804$\pm$0.002 | 0.658$\pm$0.002 | 0.838$\pm$0.005 | 0.817$\pm$0.009 | 0.820$\pm$0.006 | 0.857$\pm$0.004 | **0.887**$\pm$0.003 |
| | F1-score | 0.707$\pm$0.006 | 0.551$\pm$0.002 | 0.768$\pm$0.009 | 0.728$\pm$0.012 | 0.708$\pm$0.009 | 0.802$\pm$0.006 | **0.832**$\pm$0.005 |
| 20 | Accuracy | 0.787$\pm$0.002 | 0.655$\pm$0.001 | 0.833$\pm$0.002 | 0.812$\pm$0.001 | 0.817$\pm$0.003 | 0.845$\pm$0.002 | **0.873**$\pm$0.001 |
| | F1-score | 0.680$\pm$0.004 | 0.549$\pm$0.007 | 0.764$\pm$0.002 | 0.724$\pm$0.002 | 0.703$\pm$0.003 | 0.783$\pm$0.005 | **0.815**$\pm$0.003 |

tion network (because only the citation network has class labels) and link prediction on both citation and coauthoring network. Additionally, we demonstrate the semantic interpretability of MAI-ECS in terms of multi-aspect influences and multi-aspect interactions.

The following algorithms are compared in our paper:

- **Node2vec** (Grover and Leskovec 2016): which learns node representation through random walk and the skip-gram model with network structure only. The sampling strategy in DeepWalk (Li et al. 2014) can be seen as a special case of Node2vec.
- **Doc2vec** (Le and Mikolov 2014): which embeds text, i.e., node content, into a distributed vector by neural network models without network structure.
- **NV+DV**: which concatenates node representation from Node2vec and Doc2Vec to utilize both network structure and node content.
- **TriDNR** (Pan et al. 2016): which utilizes network, text information and labels to learn network representation.
- **CANE** (Tu et al. 2017): which is the state-of-the-art algorithm that utilizes both network and text information to learn network representation.
- **MAI-LS**: which only includes the multi-aspect interaction learning system without considering multi-aspect influence propagation.
- **MAI-ECS**: which is the full model proposed in this paper.

The node embedding length in representation methods is set to 100. The interaction and influence aspects are set to 10. In Node2vec, the random walk control parameters p = 1 and q = 1 which lead to the best performance. In TriDNR, we use different combination weight parameter which adjusting the ratio between text information and network information in different tasks because the performance is highly sensitive to this parameter (i.e., the importance ratio of text is set to 0.8 in classification and 0.2 in link prediction). Other hyper-parameters used in the above algorithms follow the recommendation or default parameters in their original source codes.

## Node Classification

We assess the effectiveness of network representations on node classification in the citation network in which the domains are regarded as class labels. The classification results in terms of accuracy and F1-score are demonstrated in Table 1. The percentages $p\%$ of training samples are set to $p = 20$ and $p = 80$.

**Importance of considering node content.** Table 1 shows that MAI-LS and MAI-ECS achieve better performance than all other methods. Node2vec which only utilizes topological information outperforms Doc2vec which only uses node content. Meanwhile combining topological and node content contributes to the node classification.

**Importance of the way to learn node content.** Although TriDNR and CANE both incorporate network and node content, their performance is worse than the concatenation of Node2vec and Doc2vec (i.e., NV+DV) on the citation network. This is because TriDNR and CANE model the network structure and node content separately by two independent objectives. This shallow combination cannot learn a comprehensive representation that embeds the information of both network structure and node content. By contrast, MAI-LS and MAI-ECS jointly learn network structure and node content through an evolutionary learning process which largely benefits the further learning tasks.

**Contribution of influence propagation system.** According to the classification results, MAI-ECS outperforms MAI-LS because MAI-ECS explores the network with multi-influence propagation system and captures more comprehensive network structure. Also, the stable multi-aspect influences are fed into the multi-aspect interaction learning system in MAI-ECS to optimize the representation learning process in turn.

## Link Prediction

Link prediction has been widely explored in network analysis, we use network representation learned on a training network which randomly samples 90% edges from the original network and test the link prediction performance on the remaining network in terms of ranking results. In particular, we use the inner-product of node representation from Node2vec, Doc2vec, NV+DV, TriDNR and CANE as the node proximity because no special design is available in these models for evaluating node proximity or edge connecting strength. In MAI-LS and MAI-ECS, the overall multi-aspect interaction scores between nodes (cf. Eq. 8) are used to quantify the connection strength. Normalized discounted cumulative gain on top-$k$ ranked nodes (nDCG@k) is used to evaluate the ranking performance for link prediction.

According to Table 2, MAI-LS and MAI-ECS significantly outperform other methods in terms of nDCG on both data sets. Link prediction is highly dependent on the representation of network structure, so the performance of Node2vec is much better than Doc2vec. Although TriDNR and CANE incorporate node contents and network structure, their performances on link prediction are quite lim-

Table 2: Link prediction performance w.r.t. nDCG on Coauthoring and Citation networks

| Models | Coauthor | | | Citation | | |
|---|---|---|---|---|---|---|
| | @1 | @5 | @10 | @1 | @5 | @10 |
| Node2vec | 0.884 | 0.921 | 0.931 | 0.716 | 0.809 | 0.829 |
| Doc2vec | 0.704 | 0.734 | 0.759 | 0.385 | 0.488 | 0.543 |
| NV+DV | 0.898 | 0.907 | 0.918 | 0.744 | 0.811 | 0.832 |
| TriDNR | 0.786 | 0.773 | 0.791 | 0.588 | 0.685 | 0.721 |
| CANE | 0.726 | 0.733 | 0.761 | 0.613 | 0.704 | 0.736 |
| MAI-LS | 0.956 | 0.961 | 0.966 | 0.805 | 0.874 | 0.889 |
| MAI-ECS | **0.965** | **0.968** | **0.972** | **0.820** | **0.885** | **0.898** |

ited, which shows the deficiency of the way that TriDNR and CANE incorporate network structure and node content, i.e., a simple linear combination of text information and network information. Another reason leading to the poor performance of CANE is that CANE emphasizes the text information and uses the convolutional neural network to capture the multiple influences within text which is not suitable for the short text, e.g., title.

MAI-LS and MAI-ECS incorporate node content and network structure in terms of multi-aspect interaction and influence evolution, which is more effective to deeply fuse the information from both sources. Moreover, MAI-LS and MAI-ECS learn to rank the edge formation through an evolutionary learning process on different masked networks, which is consistent with the goal of link prediction. In addition, MAI-ECS employs MAI-PS to disclose the multi-aspect node influences through the information propagation dynamics to refine MAI-LS learning, which makes MAI-ECS achieve the best performance.

## Semantic Interpretability Demonstration

The most distinctive characteristic of MAI-ECS is to model the multi-aspect influences and multi-aspect interactions, which empowers MAI-ECS with the semantic interpretability. We respectively list top-3 ranked authors with their research interests in two aspects according to the multi-aspect influential scores (cf. Eq. 2) output from MAI-PS. As shown in Figure 4, the research interests of the top-3 authors in aspect 1 are all about “high performance” and “systems”, and that in aspect 2 are about “network” and “communication”. These aspects reflect different research areas in the coauthoring network. Only with the text similarity between node content, we can neither differentiate the research areas because there are some commonly used words, such as “algorithm” and “model”, nor assign authors to their specific area for ranking. MAI-ECS jointly model node content, network structure, and influence propagation, which enables it to provide multi-aspect semantics and influence ranking.

To demonstrate the multi-aspect interactions between nodes, we randomly sample a paper with the title *“A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms”* and all its references from the citation network. Due to the space limitation, we only list the top-5 papers on each aspect according to the interaction scores (cf. Eq. 13) in two focused aspects 3 and 8, and this paper has few references in other aspects. As shown in Figure 5, the word clouds demonstrate the frequency of words appearing in the titles of all references. We find aspect 3 mainly focuses on “stereo” while aspect 8 mainly focuses on “tracking” and “motion”. Obviously, current state-of-the-art methods can only predict whether existing citation between papers, but they cannot semantically tell why this citation is formed and which aspect leads to this citation.

| Aspect | Top-rank Authors | Research Interests |
|---|---|---|
| 1 | Robert van de Geijn | high performance; high-performance implementation; parallel linear algebra package; parallel implementation; dense linear algebra algorithm; excellent performance |
| | Jack J. Dongarra | high performance; performance analysis; performance result; high performance computing platform; performance issue; peak performance; performance penalty |
| | Oscar H. Ibarra | P system; equivalence problem; reversal-bounded counter; neural P system; decision problem; computing power; emptiness problem; reachability problem; SN P system |
| 2 | Mario Gerla | wireless network; network coding; vehicular network; mobile network; network performance; mobile wireless network; network topology; sensor network; mesh network; |
| | Willy Yonkers | remote communication; familiar telecommunication product; human need; human product interaction; network infrastructure; physical closeness; visionary design |
| | M. Ott | future network availability; mobile device; active network monitoring; multiple access network; network availability prediction; network resource; network selection |

Figure 4: The top-3 authors on aspect 1 and aspect 2 in terms of node influence.

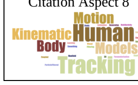
| Sample paper | *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms.* |
|---|---|
| Citation Aspect 3   | A maximum likelihood stereo algorithm.<br>Genetic-based stereo algorithm and disparity map evaluation.<br>Real-time correlation-based stereo vision with reduced border errors.<br>Calculating dense disparity maps from color stereo images.<br>Map-based stochastic diffusion for stereo matching and line fields estimation. |
| Citation Aspect 8   | Human motion tracking with a kinematic parameterization of extremal contours.<br>Learning generative models for multi-activity body pose estimation.<br>A study on smoothing for particle-filtered 3d human body tracking.<br>Optimization and filtering for human motion capture.<br>Coupled visual and kinematic manifold models for tracking. |

Figure 5: The top-5 papers on aspect 3 and aspect 8 in terms of the interaction scores.

# Conclusion and Future Work

In this paper, we propose a critical and challenging problem in attributed network representation: learning the heterogeneities of multi-aspect interactions between nodes and node influences in a complex network. Accordingly, we propose the MAI-ECS to model multi-aspect node influence and multi-aspect interactions between nodes. MAI-ECS enables an evolutionary learning process on integrating node texts and network structure to jointly and iteratively learn the multi-aspect influences and interactions.

There are many future research directions as MAI-ECS is flexible to learn weighted, directed, and attributed network. First, MAI-ECS can be extended to handle dynamical networks by modifying the evolutionary learning process to a time-dependent state model. Second, MAI-ECS can be customized for community detection and role detection because the multi-aspect influences can be regarded as a node’s community or roles in a network. Also, MAI-ECS has potential to infer the semantics of node text, such as topic deviation and text mining, as shown in the demonstration.

## Acknowledgement

This work is partially supported by the The National Key Research and Development Program of China (2016YFB0200401), by the program for New Century Excellent Talents in University, by National High-level Personnel for Defense Technology Program (2017-JCJQ-ZQ-013), by the HUNAN Province Science Foundation 2017RS3045.

## References

[Cai, Zheng, and Chang 2017] Cai, H.; Zheng, V. W.; and Chang, K. C.-C. 2017. A comprehensive survey of graph embedding: Problems, techniques and applications. *arXiv preprint arXiv:1709.07604*.

[Cao, Lu, and Xu 2015] Cao, S.; Lu, W.; and Xu, Q. 2015. Grarep: Learning graph representations with global structural information. In *CIKM*, 891–900. ACM.

[Cao 2014] Cao, L. 2014. Non-iidness learning in behavioral and social data. *The Computer Journal* 57(9):1358–1370.

[Cao 2015] Cao, L. 2015. Coupling learning of complex interactions. *Information Processing & Management* 51(2):167–186.

[Cui et al. 2018] Cui, P.; Wang, X.; Pei, J.; and Zhu, W. 2018. A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering*.

[Gao and Huang 2018] Gao, H., and Huang, H. 2018. Deep attributed network embedding. In *IJCAI*, 3364–3370.

[Geiser 2014] Geiser, J. 2014. *Coupled Systems: Theory, Models, and Applications in Engineering*. Chapman and Hall/CRC.

[Grover and Leskovec 2016] Grover, A., and Leskovec, J. 2016. node2vec: Scalable feature learning for networks. In *SIGKDD*, 855–864. ACM.

[He and Niyogi 2004] He, X., and Niyogi, P. 2004. Locality preserving projections. In *Advances in neural information processing systems*, 153–160.

[Hu et al. 2018] Hu, L.; Jian, S.; Cao, L.; and Chen, Q. 2018. Interpretable recommendation via attraction modeling: Learning multilevel attractiveness over multimodal movie contents. In *IJCAI*, 3400–3406.

[Hu et al. 2019] Hu, L.; Jian, S.; Cao, L.; Gu, Z.; Chen, Q.; and Amirbekyan, A. 2019. Hers: Modeling influential contexts with heterogeneous relations for sparse and cold-start recommendation. In *AAAI*.

[Jang, Gu, and Poole 2016] Jang, E.; Gu, S.; and Poole, B. 2016. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*.

[Le and Mikolov 2014] Le, Q., and Mikolov, T. 2014. Distributed representations of sentences and documents. In *International Conference on Machine Learning*, 1188–1196.

[Li et al. 2014] Li, X.; Du, N.; Li, H.; Li, K.; Gao, J.; and Zhang, A. 2014. A deep learning approach to link prediction in dynamic networks. In *SIAM*, 289–297.

[Mikolov et al. 2013] Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

[Pan et al. 2016] Pan, S.; Wu, J.; Zhu, X.; Zhang, C.; and Wang, Y. 2016. Tri-party deep network representation. *Network* 11(9):12.

[Pennington, Socher, and Manning 2014] Pennington, J.; Socher, R.; and Manning, C. 2014. Glove: Global vectors for word representation. In *EMNLP*, 1532–1543.

[Perozzi, Al-Rfou, and Skiena 2014] Perozzi, B.; Al-Rfou, R.; and Skiena, S. 2014. Deepwalk: Online learning of social representations. In *SIGKDD*, 701–710. ACM.

[Ruder 2016] Ruder, S. 2016. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

[Saad 2011] Saad, Y. 2011. *Numerical methods for large eigenvalue problems: revised edition*, volume 66. SIAM.

[Simonyan and Zisserman 2014] Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[Tang et al. 2015] Tang, J.; Qu, M.; Wang, M.; Zhang, M.; Yan, J.; and Mei, Q. 2015. Line: Large-scale information network embedding. In *WWW*, 1067–1077. International World Wide Web Conferences Steering Committee.

[Tu et al. 2017] Tu, C.; Liu, H.; Liu, Z.; and Sun, M. 2017. Cane: Context-aware network embedding for relation modeling. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, 1722–1731.

[Wang et al. 2017] Wang, X.; Cui, P.; Wang, J.; Pei, J.; Zhu, W.; and Yang, S. 2017. Community preserving network embedding. In *AAAI*, 203–209.

[Wang, Cui, and Zhu 2016] Wang, D.; Cui, P.; and Zhu, W. 2016. Structural deep network embedding. In *SIGKDD*, 1225–1234. ACM.

[Yang et al. 2015] Yang, C.; Liu, Z.; Zhao, D.; Sun, M.; and Chang, E. Y. 2015. Network representation learning with rich text information. In *IJCAI*, 2111–2117.

[Yin, Gao, and Lin 2016] Yin, M.; Gao, J.; and Lin, Z. 2016. Laplacian regularized low-rank representation and its applications. *IEEE transactions on pattern analysis and machine intelligence* 38(3):504–517.

[Zhang et al. 2017] Zhang, D.; Yin, J.; Zhu, X.; and Zhang, C. 2017. User profile preserving social network embedding. In *Proceedings of IJCAI*, 3378–3384.

[Zhang et al. 2018a] Zhang, Q.; Cao, L.; Zhu, C.; Li, Z.; and Sun, J. 2018a. Coupledcf: Learning explicit and implicit user-item couplings in recommendation for deep collaborative filtering. In *IJCAI*, 3662–3668.

[Zhang et al. 2018b] Zhang, Z.; Yang, H.; Bu, J.; Zhou, S.; Yu, P.; Zhang, J.; Ester, M.; and Wang, C. 2018b. Anrl: Attributed network representation learning via deep neural networks. In *IJCAI*, 3155–3161.