Identifying High Potential Talent: A Neural Network based Dynamic Social Profiling Approach

Yuyang Ye[1,2], Hengshu Zhu[2,*], Tong Xu[1], Fuzhen Zhuang[3,4], Runlong Yu[1], Hui Xiong[1,2,*]
[1]Anhui Province Key Lab of Big Data Analysis and Application, University of Science and Technology of China
[2]Baidu Talent Intelligence Center, Baidu Inc
[3]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology
[4]University of Chinese Academy of Sciences
{yeyuyang, yrunl}@mail.ustc.edu.cn, zhuhengshu@baidu.com, tongxu@ustc.edu.cn,
zhuangfuzhen@ict.ac.cn, xionghui@gmail.com

***Abstract*—How to identify high-potential talent (HIPO) earlier in their career always has strategic importance for human resource management. While tremendous efforts have been made in this direction, most existing approaches are still based on the subjective selection of human resource experts. This could lead to unintentional bias and inconsistencies. To this end, in this paper, we propose a neural network based dynamic social profiling approach for quantitatively identifying HIPOs from the newly-enrolled employees by modeling the dynamics of their behaviors in organizational social networks. A basic assumption is that HIPOs usually perform more actively and have higher competencies than their peers to accumulate their social capitals during their daily work practice. Along this line, we first propose to model the social profiles of employees with both Graph Convolutional Network (GCN) and social centrality analysis in a comprehensive way. Then, an adaptive Long Short Term Memory (LSTM) network with global attention mechanism is designed to capture the profile dynamics of employees in the organizational social networks during their early career. Finally, extensive experiments on real-world data clearly validate the effectiveness of our approach as well as the interpretability of our results.**

***Index Terms*—Human Resource Management, HIPO identification, Social Profiling**

I. INTRODUCTION

In the current fast-evolving business environment, modern companies are under pressure to constantly improve their talent selection and development strategies for maintaining the competitive edge and supporting their long-term business development goals. High potential talents (HIPOs) are often regarded as future leaders within organizations. Compared to their peers, they have the leadership ability, business acumen, and the desire for success and usually advance at a faster pace [1]. Given the significant role of HIPOs in the execution of organizational strategy and the optimization of organizational structure [2], it is always a major concern in human resource management to prospectively identify and develop HIPOs [3], especially among newly-enrolled employees, so that special attention could be paid to cultivate the future leaders [4].

In the past decades, while tremendous efforts have been made in this direction, existing approaches are usually based on the subjective selection of human resource experts. This could lead to unintentional bias and inconsistencies [5]. Alternatively, with the rapid development of management information systems, large-scale talent data have been accumulated and provide an unparalleled opportunity for business leaders to identify talents in a data-driven way. To this end, in this paper, we propose to develop a data-driven approach for quantitatively identifying HIPOs from the newly-enrolled employees by modeling the dynamics of their behaviors in organizational social networks. As shown in Figure 1, a basic assumption is that HIPOs usually perform more actively and have higher competencies than their peers to accumulate their social capitals during their daily work practice [6]–[10]. Indeed, when we may not be able to identify HIPOs from their peers by some explicit cues (e.g. job title or salary), we could still detect HIPOs through other implicit information sources [11]: such as their social ties and the centrality of talents in the organizational social networks. Thus we aim to utilize machine learning techniques to quantitatively address the HIPO identification problem and introduce the additional insight from the social profiling perspective.

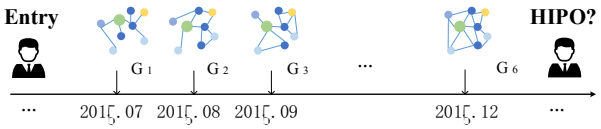


Fig. 1: A motivating example of HIPO identification.

Along this line, there are still two major challenges to be addressed. First, in order to effectively represent the employees' social capitals, it is necessary to model the profiles of employees in the organizational social networks in a comprehensive way. Second, how to dynamically capture the change of the social profiles of employees, and meanwhile, provide interpretability for domain experts in HIPO identification. Therefore, in this paper, we propose a novel neural network based dynamic social profiling (NNDSP) approach for early identifying HIPOs. Specifically, we first propose to comprehensively model the social profiles of employees with both Graph Convolutional Network (GCN) and social

* denotes the corresponding author

2374-8486/19/$31.00 ©2019 IEEE
DOI 10.1109/ICDM.2019.00082

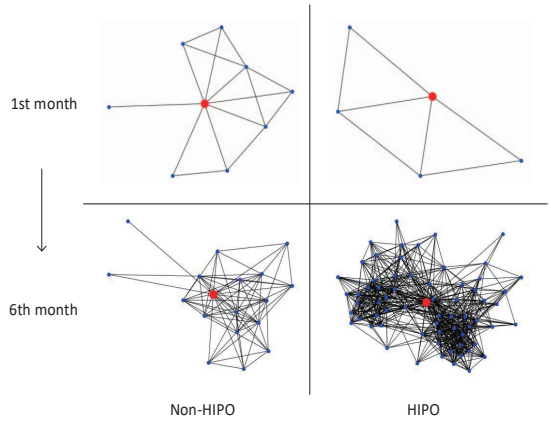


Fig. 2: An example of social network evolution between HIPO and Non-HIPO.

centrality analysis. In particular, for each employee, we extract both local and global information in the organizational social networks as social profiles, and integrate them with the basic profile of employees as features for HIPO identification. Then, an adaptive Long Short Term Memory (LSTM) network with global attention mechanism is designed to capture the profile dynamics of the employees in the organizational social networks during their early career. Finally, we evaluate our approach with extensive experiments on real-world talent data. The experimental results clearly validate the effectiveness and interpretability of NNDSP for HIPO identification.

Overview. The rest of this paper is organized as follows. In Section II, we introduce the preliminaries of this paper and formulate the problem of HIPO identification. In Section III, we introduce the technical details of our approach NNDSP. In Section IV, we comprehensively evaluate the performances of our model on a real-world dataset. We briefly introduce the related works of this paper in Section V. Finally, in Section VI, we conclude the paper.

II. Preliminaries

In this paper, we propose to develop a data-driven approach for quantitatively identifying HIPOs from the newly-enrolled employees, through modeling the dynamics of their organizational social networks.

The real-world talent data used in this work, which consist of two datasets, were provided by a high-tech company in China. The first is the monthly organizational social networks, which were collected from the internal email system from January 2015 to January 2018. Specifically, each network can be represented as an undirected graph $G_j = (V_j, E_j)$, where V_j is the set of employees who have email communications in the j-th month, and $e_{ij} \in E_j$ means employees $v_i, v_j \in V_j$ have at least one email communication in this month. In this dataset, for each employee v_i, we also have some job-related information, such as job position, job level and reporting lines, which can be used as basic features in our HIPO identification model. The second dataset is the HIPO list, which records the employees who had been named as HIPOs in the first year after they joined the company. In particular, to avoid the impact of external work experiences, in this paper we only studied the employees who were enrolled from campus recruitment with the position of engineer. Note that, all the sensitive information in the datasets (e.g., name) has been anonymized for the privacy-protection purpose.

Based on the above talent data, in this paper we aim to predict whether an employee could be selected as a HIPO in the first year after she joined the company, with the observation of her organizational social network. For preliminarily validating our assumption, we randomly selected a pair of employees labeled as HIPO and Non-HIPO respectively in our dataset, and drew their one-hop social networks in the first and sixth month after they joined the company. As shown in Figure 2, we can find that the sparsities of two social networks are similar in the first month. However, after a few months, the HIPO's social network becomes much denser than that of the Non-HIPO. This result clearly validates our motivation of modeling the social network dynamics for HIPO identification. To be specific, the problem of HIPO identification in this paper can be formally defined as follows.

Problem Definition. Given a new employee v_i, who joined the company in the t_i-th month, and k organizational social networks $\mathcal{G}_i = \{G_{t_i}, \cdots, G_{t_i+k}\}$, the objective is to develop a model $f(v_i, \mathcal{G}_i) = y$ to predict whether v_i is a HIPO (i.e., $y = 1$) or not (i.e., $y = 0$).

Note that, to achieve the early identification of HIPOs, in this paper, we define $k \leq 6$, which equals to the probation period of new employees in the company. In the following section, we will introduce the technical details of our neural network based approach for addressing the proposed HIPO identification problem.

III. Technical Details

In this section, we will introduce the technical details of our NNDSP for HIPO identification. Specifically, the network structure of NNDSP is shown in Figure 3, which mainly consists of three components, namely employee social profiling, social dynamic modeling and HIPO identification.

A. Employee Social Profiling

A basic assumption of our approach is that HIPOs usually perform more actively and have higher competencies to accumulate their social capitals during their work practice, compared with other common employees. Therefore, in NNDSP, we first propose to comprehensively model the social profiles of employees, for effectively representing their social capitals. To this end, we define two types of information in the organizational social networks as the social profiles of employees, namely local social information and global social information. Specifically, the local social information depicts the social contexts formed by the close colleagues of employees, indicating the direct social ties in the daily work activities (e.g., projects). Besides, the global social information reflects

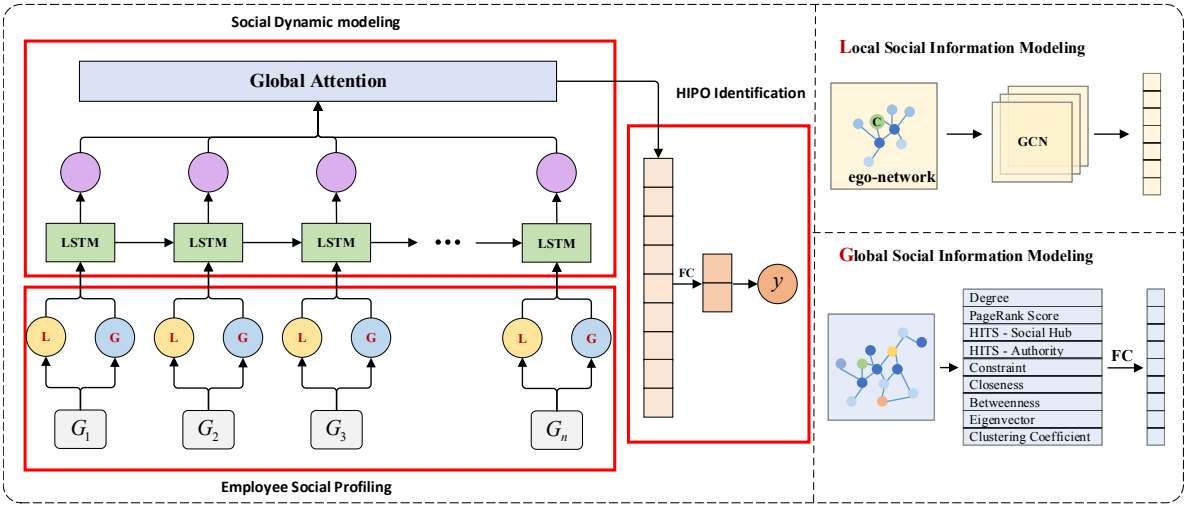


Fig. 3: An illustration of NNDSP for HIPO Identification.

the social status of employees in the whole organizational social networks, indicating whether they played a significant role in the company (e.g., structure hole).

1) Local Social Information Modeling: Here we propose to use Graph Convolutional Network (GCN) for modeling the local social information of employees. Specifically, given an organizational social network G, we define $A \in \mathbb{R}^{n\times n}$ as its adjacency matrix and D as the degree matrix, where $D_{ii} = \sum_j A_{ij}$. Meanwhile, $X \in \mathbb{R}^{n\times m}$ is the input feature matrix of employees, where each row is an m-dimension feature representation of employee $v \in V$. To model the local social information with GCN, we first calculate $\hat{A} = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$, where $\tilde{A} = A+I_N$, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$, and I_N is the identity matrix. Then, after adapting first-order approximation of localized filters, the representation matrix S of G can be calculated as:

$$S = f(X, A) = \alpha(\hat{A}XW + b), \tag{1}$$

where α is the activation function, and $W \in \mathbb{R}^{d\times m}$ and $b \in \mathbb{R}^d$ are parameters to be learned in the model training stage.

As mentioned in Section 2, for each new employee v_i, we got a set of networks $\mathcal{G}_i$ which has k elements. Each network G_j in this set denotes the organizational social network in the j-th month. In order to capture the social context of employees and eliminate the influence from the irrelevant colleagues, we build a type of sub-graph, called ego-network (only contains a central node and its neighbors), based on each network in $\mathcal{G}_i$. The central node of each ego-network is v_i. And each ego-network contains the first-order neighbors and the second-order neighbors of the central node, which respectively denote the nodes which have direct interaction and indirect interaction with the central node in the organizational networks. Thus, for new employee v_i, we have a set of ego-networks $\{G^e_{ij}\}_{j=1}^{k}$. In particular, we let A^i_j denote the adjacency matrix of G^e_{ij}. By using feature embedding operation on the time-vary factors, we also get the feature matrix $X_j \in \mathbb{R}^{n\times m}$ whose rows are the feature vectors of the employees existing in the organizational networks of the j-th month. Particularly for capturing influence from the second-order neighbors, we propose a two-layer GCN to get the representation matrix $S_{ij} \in \mathbb{R}^{n\times d}$ of G^e_{ij} as:

$$S_{ij} = f(X_j, A^i_j) = \alpha(\hat{A}^i_j\ \alpha(\hat{A}^i_j X_j W_0 + b_0)W_1 + b_1), \tag{2}$$

where W_0, b_0 and W_1, b_1 are parameters of two layers respectively. Each row in the representation matrix S_{ij} denotes the representation of a node in G^e_{ij}.

Thus, by choosing the corresponding row of the central node v_i, we can get the vector s^l_{ij} to represent the social contextual information of v_i in the organizational networks in the j-th month.

2) Global Social Information Modeling: Other than local social information, the social status in the whole organizational social networks is also important for profiling the social capitals of employees, which are defined as the global social information in NNDSP. Specifically, according to the state-of-the-art studies in social network analysis [12], [13], here we propose to use the social centrality scores for modeling the global social information of employees.

Specifically, given a graph $G = (V, E)$, where the number of vertexes is N, we propose to use 9 different social centrality factors in our experiments, which are defined as follows.

- **PageRank.** PageRank is an algorithm used by Google Search to rank web pages, which can measure the importance of nodes in a graph. After sharing initialized PageRank score to each node as $1/N$, the PageRank score of node u can be calculated as:

$$PR(u) = d(\sum_{v\in N_u} \frac{PR(v)}{L(v)}) + (1-d)/N, \tag{3}$$

where N_u is the neighbors of node n, $L(v)$ is the number of edges whose origin node is v, d is the damping factor. After several iterations until convergence, the PageRank scores of all nodes can be obtained.

- **Hubs and Authorities.** Hubs and Authorities are calculated by HITS algorithm [14], which can estimate the value of the links from a node and the value of node itself respectively. Like PageRank algorithm, we set the $hub(u) = 1$ and $auth(u) = 1$, $\forall u \in V$ as initialization. The update process is defined as,

$$\text{auth}(u) = \sum_{i=1}^{n_1} \text{hub}(i), \ \text{hub}(u) = \sum_{j=1}^{n_2} \text{auth}(j), \quad (4)$$

where n_1 is indegree of node u, n_2 is outdegree of node u, i is a node which has an edge linked to u, j is a node which has an edge linked from u. After several iterations of updates, we can get the final Hubs and Authorities of all nodes.

- **Constraint.** The concept of node constraint is proposed by [15]. Node constraint is a measurement in Structural Hole Theory. The node constraint of node $u \in V$ is defined as:

$$c_u = \sum_{u \neq v} \left(p_{uv} + \sum_{w \neq u, w \neq v} p_{uw} p_{vw} \right)^2, \quad (5)$$

where p_{ij} is $1/degree(i)$ if e_{ij} in E else $p_{ij} = 0$. The network constraint measures the extent to which a person's contacts are redundant. The lower the score on the node constraint is, the more essential position a node in the information flow between the two nodes occupies.

- **Degree centrality.** Degree centrality of a node $u \in V$ is just defined as the number of edges u has, i.e.,

$$C_D(u) = \text{degree}(v). \quad (6)$$

- **Closeness centrality.** The closeness centrality of a node is defined as the average length of the shortest path between the node and all other nodes in the graph [16]. The closeness centrality of node $u \in V$ can be calculated as follows:

$$C_C(u) = \frac{1}{\sum_{v \in V} d(v, u)}, \quad (7)$$

where $d(u, v)$ denotes the distance between node u and node v.

- **Betweenness centrality.** The betweenness centrality is measure introduced by [17] that quantifies the number of times a node acts as a essential itermediate node in the shortest path between two other nodes. The betweenness centrality of node $u \in V$ is defined as:

$$C_B(u) = \sum_{v \neq u \neq w \in V} \frac{\sigma_{vw}(u)}{\sigma_{vw}}, \quad (8)$$

where σ_{vw} is the number of shortest path between v and w and $\sigma_{vw}(u)$ denotes the number of these paths which go through u.

- **Eigenvector centrality.** The eigenvector centrality is a measure of the influence of a node in the graph. A high eigenvector score means that a node is connected to many nodes who themselves have high scores [18]. Let $A = (a_{i,j})$ be the adjacency matrix of G. The eigenvector centrality x_u of node $u \in V$ is defined as:

$$x_u = \frac{1}{\lambda} \sum_v a_{v,u} \, x_v, \quad (9)$$

where $\lambda \neq 0$ is a constant. In matrix form we have $\lambda x = xA$. The update process can be formulated as:

$$\begin{aligned} x^{(k)} &= x^{(k-1)} A, \\ x^{(k)} &= x^{(k)} / max(x^{(k)}), \end{aligned} \quad (10)$$

where $x^{(k)}$ denotes the eigenvector of A after k-step update. The rate of convergence is the rate at which $(\lambda_2/\lambda_1)^k$ goes to 0, while λ_1 and λ_2 respectively denote the largest and the second largest eigenvalues of A.

- **Clustring coefficient.** The clustring coefficient ia a measure that quantifies how close the neighbors of a node in the graph be a complete graph [19]. Let N_u be the neighbors of $u \in G$ in G and the number of elements in N_u is k_u, the clustring coefficient of u is given as:

$$C_u = \frac{2 \left| \{ e_{vw} : v, w \in N_u, e_{vw} \in E \} \right|}{k_u \left(k_u - 1 \right)} \quad (11)$$

Based on our real-world dataset, Figure 4 shows the distribution of these 9 social centrality factors in each month since employees entered the company. Intuitively, we can find that both HIPOs and Non-HIPOs can promote their social centrality in the organizational networks, i.e., accumulated their social capitals. Moreover, compared with Non-HIPOs, HIPOs can promote their social centrality factors more effectively in terms of both speed and numerical value.

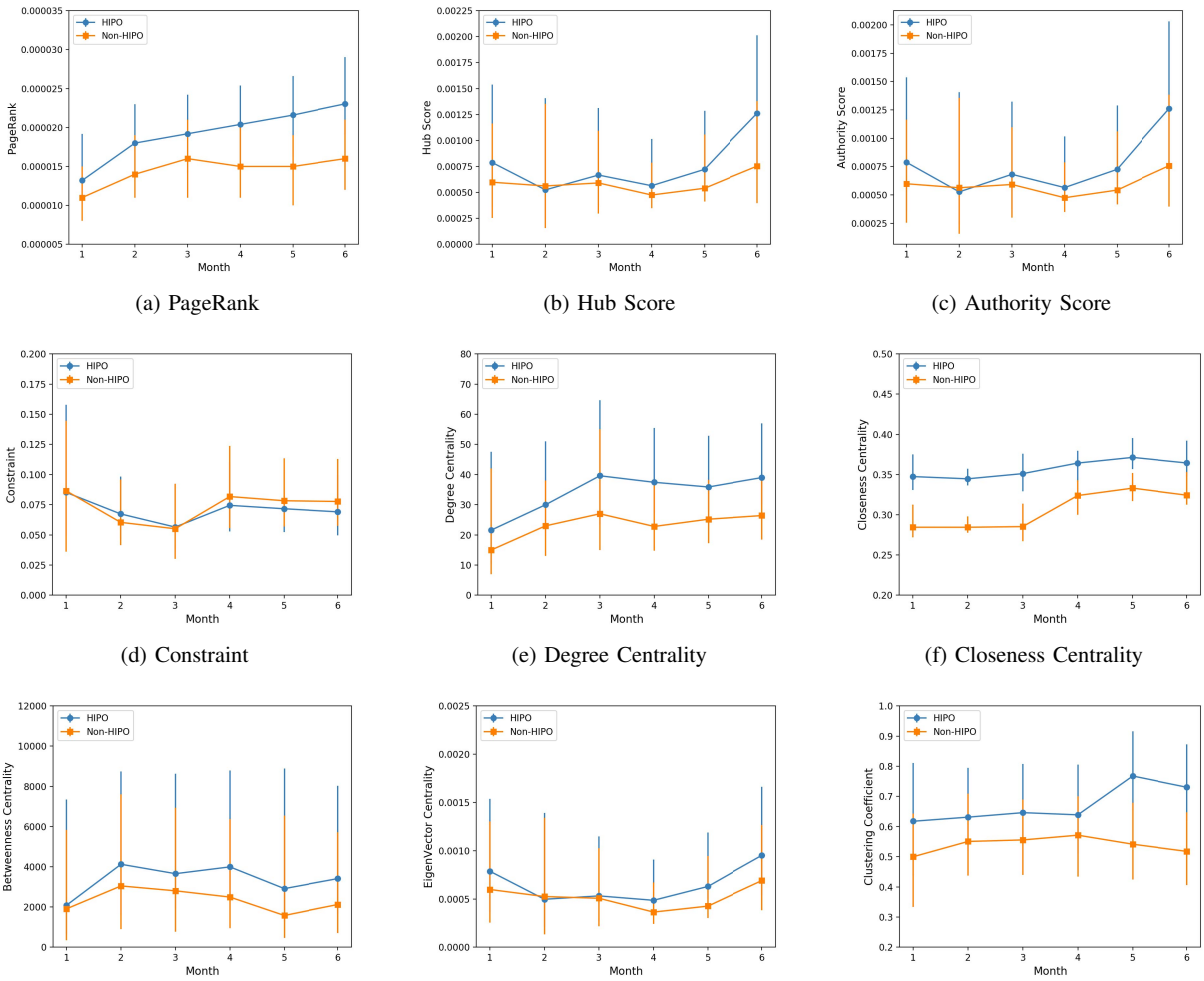
Therefore, in each month, we can obtain a 9-dimension vector consists of the normalized social centrality scores for each employee. After that, we propose to use a fully-connected embedding layer to transform this vector into a latent representation s_{ij}^g as the output of global social information modeling for employee v_i in the j-th month.

Based on the local social information representation s_{ij}^l and the global information representation s_{ij}^g, we can use the concatenation operation to obtain the final social profile representation r_j^i of employee v_i in the organizational social network G_j:

$$r_j^i = [s_{ij}^l; s_{ij}^g]. \quad (12)$$

Furthermore, we can get the sequence of representation $\left\{ r_{t_i}^i, \cdots, r_{t_i+k}^i \right\}$ for the employee v_i who joined the company in the t_i-th month.

3) Feature Representation: Here, we introduce the generation of the feature representation matrix X in the GCN for local social profile modeling, which is generated by embedding the employee features into a representation space. Specifically, some major features used in this paper are listed in the Table I.



(a) PageRank (b) Hub Score (c) Authority Score

(d) Constraint (e) Degree Centrality (f) Closeness Centrality

(g) Betweenness Centrality (h) Eigenvector Centrality (i) Clustring Coefficient

Fig. 4: The distribution of social centrality factors in each month since employees entered the company.

For each numerical feature, we adopt a normalization operation, in order to prevent the prediction being dominated by the factor which has a large range of distribution. Meanwhile, for the categorical features, we utilize the one-hot encoder as representation. In addition, for the missing value, we adopt the average pooling operation. Finally, all the features are concatenated into a representation vector and fed into a fully-connected embedding layer to obtain a 20-dimension feature vector X_i for each employee $v_i \in V$.

B. Network Dynamic Modeling

After social profiling, we can obtain a sequence of representation vectors for each employee. Then, we propose to leverage the Long Short Term Memory (LSTM) network [20], which can store and access a long range of contextual information in the sequential input, to capture the profile dynamics of employees in the organizational social networks. Specifically, the LSTM network has a cell state and three gates, named as the input gate i, forget gate f and output gate o, which can be formulated as follows:

$$\begin{aligned} &i_t = \sigma\left(W_i\left[r_t, h_{t-1}\right] + b_i\right), f_t = \sigma\left(W_f\left[r_t, h_{t-1}\right] + b_f\right), \\ &C_t = f_t \odot C_{t-1} + i_t \odot \tanh\left(W_C\left[r_t, h_{t-1}\right] + b_C\right), \\ &o_t = \sigma\left(W_o\left[r_t, h_{t-1}\right] + b_o\right), h_t = o_t \odot \tanh\left(C_t\right) \end{aligned} \tag{13}$$

where $X = \{r_1, \cdots, r_k\}$ is the input sequence. And W_f , W_i, W_C, W_o, b_f, b_i, b_C, b_o are the parameters to be trained in the training process, $\odot$ represents element wise multiplication, σ is the sigmoid function, and $\{h_1, h_2, \cdots, h_k\}$ represents a sequence of hidden representation in the LSTM.

Furthermore, in order to improve the interpretability of our model, we propose to use a global attention mechanism [21] to quantitatively evaluate the contributions of employee's social profiles in each month to the result of HIPO identification. Specifically, we set $\left\{h_i^1, \cdots, h_i^k\right\}$ as the sequential output of the LSTM for the input sequence of employee v_i, the attention

TABLE I: Some descriptions of employee features.

Type	Description
Numerical	The frequency of communication with subordinates and superiors; The number of subordinates and superiors; The average and variance of working hours per day; Age; The length of service;
Categorical	Job level and position; Job department; Educational background;Gender

layer can be formulated as:

$$\begin{aligned} e_i^j &= \tanh\left(W_i h_i^j + b_i\right), \\ \alpha_i^j &= \frac{\exp\left(e_i^j\right)^\top v_\alpha}{\sum_{t=1}^{k} \exp\left(e_i^t\right)^\top v_\alpha}, \\ z_i &= \sum_{t=1}^{k} \alpha_i^t \left(W_i h_i^t + b_i\right), \end{aligned} \tag{14}$$

where α_i^j is the normalized weight for the i-th employee's information in the j-th month, z_i is the final output representation. $\{W_i\}_i$, $\{b_i\}_i$ and v_α are parameters to be learned in the training process.

C. HIPO Identification

With the learned representation of each employee through the LSTM network with global attention mechanism, we can formulate the final task of HIPO identification as a binary classification problem. Therefore, in the final component of NNDSP, we use a fully-connected layer with softmax function to learn a 2-dimension vector d for predicting the label y, which can be formulated as follows:

$$\begin{aligned} d &= W_d z + b_d, \\ y &= \mathrm{Softmax}(W_y d + b_y), \end{aligned} \tag{15}$$

where W_d, b_d, W_y , b_y are the parameters to be learned, and $y \in [0, 1]$ is the final output of our model. If $y \geq 0.5$, the predictive label of this employee is 1, which means the employee is a HIPO, and vice versa. Specifically, here we minimize the binary cross entropy for training NNDSP.

IV. Experiments

In this section, we evaluate our NNDSP model with extensive experiments on real-world talent data. We will give some discussion on the experimental results. Meanwhile, we also analyze the attention weights learnt by our model in case study, and introduce some case studies.

A. Experimental Setup

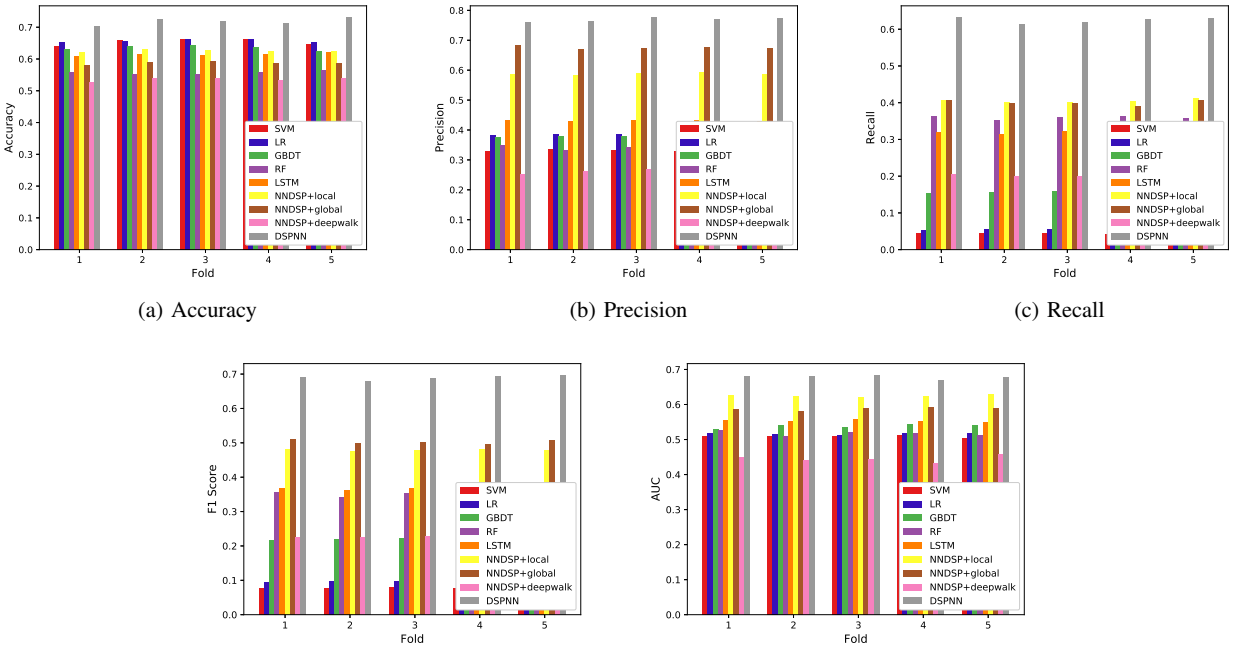
- **Data Preprocessing.** As introduced in Section II, the dataset used in our experiments were provided by a high-tech company in China. In the dataset, there are 36 month-level organizational social networks from January 2015 to January 2018, including totally 95,919 unique employees and each of them has 31 profile attributes. In each month, if two employees have interactions with each other, there would be an edge between them in the corresponding month-level network. In the experiments, to avoid the impact of external work experiences, we only predicted the employees who were enrolled from campus recruitment with the position of engineer, and removed the employees who stayed in the company for less than 6 months. After that, we totally obtained 1,341 employees as candidates for HIPO identification. As mentioned before, we have the HIPO list, which records the employees who had been named as HIPOs in the first year after they joined the company and can be used as the ground-truth. Specifically, in the experiments, there are 421 positive samples (i.e., HIPOs) and 920 negative samples for evaluation. Furthermore, our experiments were conducted with five-fold cross validation where we randomly selected 70% samples as training data, 10% as validation data, and the rest 20% as test data.
- **Parameters.** In NNDSP, we set the dimension of hidden embedding as 128 and 64 in the first and the second layers of the 2-layer GCNs, respectively. Meanwhile, in the LSTM, we set the dimension of the hidden state as 128 to capture the dynamic change of employee's social profile in organizational networks. Moreover, in the training process, we set the batch size as 64 and used the dropout layer with the probability as 0.5 to prevent the overfitting. Besides, we adopted the undersampling method in the model learning process. Finally, we used the Adam algorithm to optimize our model where the learning rate α was set to 0.01.
- **Evaluation Metrics** Indeed, the HIPO identification task in this paper can be regarded as a binary classification problem. Therefore, in the experiments, we adopted the *Accuracy*, *Precision*, *Recall*, *F1-score* and *AUC* as the evaluation metrics to validate the performance of NNDSP and other baselines.

B. Benchmark Methods

We compared NNDSP with several state-of-the-art baselines to validate its effectiveness. Specifically, baselines can be listed as follows:

- *SVM* (with linear kernel), *Logistic Regression*, which are two classic supervised models for classification task. We used the concatenation of employee features of each month as input.
- *Gradient Boosting Decision Tree* [22], *Random Forests* [23], which are two representative methods with ensemble learning and perform well in real-world classification task. We also used the concatenation of employee features of each month as input.
- *LSTM* [20], a popular neural network models for sequential data modeling, where we used the dynamic features of the employee as input.

Moreover, we selected some adopted models for evaluating the effectiveness of some components in NNDSP, including *NNDSP+local* (i.e., only considering the local information in



(a) Accuracy (b) Precision (c) Recall

(d) F1 Score (e) AUC

Fig. 5: The overall performance of NNDSP and baselines in terms of HIPO identification under five-fold cross validation.

the social profile modeling), *NNDSP+global* (only considering the global information in the social profile modeling) and *NNDSP+DeepWalk* (using DeepWalk [24], a network embedding approach which aims to get node representations in the network, instead of GCN for local information modeling). Note that, all the parameters in the baselines were carefully tuned for a fair comparison.

C. Overall Performance

We first evaluate the overall performance of NNDSP and other baselines in terms of HIPO identification. In the experiments, we utilize a five-fold cross-validation to evaluate each test approach. To be specific, we first randomly divide the data into five equal parts, and then we use each part as the test data while using the other four parts as the training data in five test rounds. As shown in Figure 5, we can have some observations. First, we can observe that NNDSP outperforms all baselines with overall evaluation metrics. Second, among classic classification models, the linear models (i.e., Logistic Regression and SVM with the linear kernel) have relatively higher accuracy but lower F1 score than non-linear models (i.e., GBDT and Random Forests). Indeed, it is because that most of the samples are predicted as non-HIPOs by linear models, which do not have sufficient ability for distinguishing HIPOs and non-HIPOs with basic features. Third, the sequential models (i.e., LSTM and other adapted models) generally have better performance than classic supervised models which are not designed for sequential data. Indeed, this finding also validates our assumption that we should focus on the dynamic change of social profiles but not just the basic features of employees in HIPO identification. Fourth, compared with other adapted models, the LSTM has limited performance, which indicates that the social profile of employees is very important for HIPO identification. Fifth, by comparing the performance of NNDSP with NNDSP+global and NNDSP+local, we can find that both of the local information and the global information has the positive impact on the performance of HIPO identification, which validates the effectiveness of social profile modeling in NNDSP. Last, it is obvious that DeepWalk+LSTM cannot achieve a satisfied performance, which is because that classic unsupervised network embedding approach, which does not incorporate the node attributes, cannot fit the HIPO identification task and even brings negative influence on the performance. Indeed, it also validates the effectiveness of using GCN framework for extracting the local social profile of employee in organizational social networks.

D. Robustness Analysis

To validate the robustness of our NNDSP model, here we also evaluate the performance of NNDSP under the different set of training/test split, different length of observation periods (i.e., a different number of observed organizational social networks) and different training/test split strategies. Specifically, in the first evaluation, we randomly selected 80%, 70%, 60%, and 50% of data for training, 10 percent for validation, and the rest for the test. From the experimental

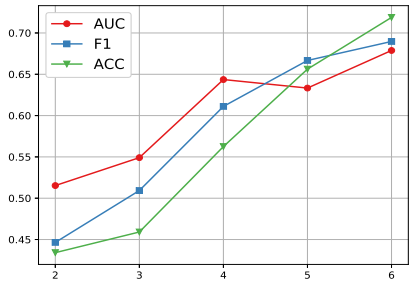


Fig. 6: Performance of NNDSP under different observations.

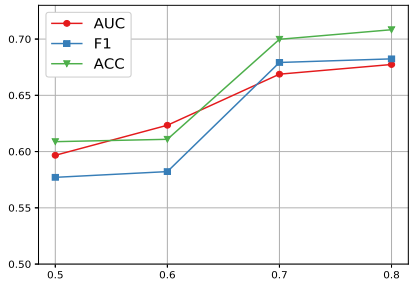
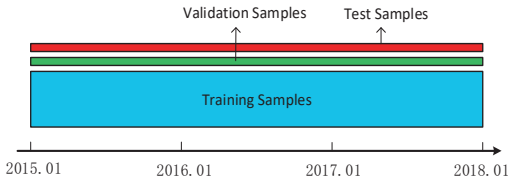


Fig. 7: Performance of NNDSP under different training ratio.

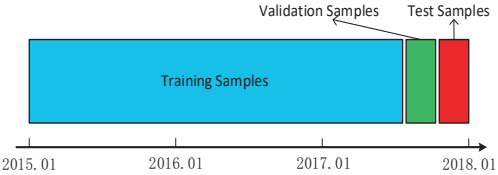
results shown in Figure 7, we can observe that, with the increase of the ratio of training data, the performance of our model will also increase to a relatively stable state. Furthermore, in the second evaluation, we set the observation periods as 6, 5, 4, 3, 2 months respectively and analyzes the performance of our model, as shown in Figure 6. We can find that NNDSP cannot perform well with limited observation periods (i.e., with less than 4 months), while it performs better with the increase of observation periods. Indeed, this result is reasonable, since it is intuitively hard to identify HIPOs with very short observation periods. In the third evaluation, we split our dataset by considering the temporal correlation, instead of randomly splitting, which uses the former 70 percent of samples as training samples and others as validation/test dataset. Figure 8 shows the examples of different strategies. Then we compare the performance with the experiment under the original random split strategy, shown in Table II. The result demonstrates the performance of NNDSP can archive a similar level and will not be influenced by different training/test split strategies. Thus our model can be used to predict the newly-enrolled employees by analyzing historical data, which is in accordance with the practical application scenarios.

TABLE II: The performance under different split strategies

Metric	ACC	PRE	REC	F1	AUC
Temporal	0.7053	0.7712	0.6074	0.6796	0.6649
Random	0.7188	0.7692	0.6250	0.6897	0.6768



(a) Randomly Split



(b) Temporally Split

Fig. 8: Examples of different split strategies.

E. Case Study

With the attention mechanism in NNDSP, we can easily obtain the contributes of the social profiles in each month to the final result of HIPO identification. To this end, here we randomly selected ten employees in the dataset and visualized the global weights for each sample that were learned by the attention mechanism, as shown in Figure 9 where redder color means higher numerical value. From the results, we can find that, for most samples the social profiles in the fourth month and the fifth month after employees joined the company make the most contributions to the HIPO identification task, which reveals that the performance of these months play a decisive role for judging whether a newly-enrolled employee would be a HIPO. Therefore, in real-world applications, human resource experts should pay more attention to the performance of employees in corresponding periods, in order to identify HIPOs in their early career. Moreover, this finding is also consistent with the results of the second robustness test on the above. Meanwhile, to further validate the effectiveness of NNDSP in terms of HIPO identification, we propose to study the latent HIPO representations learned by NNDSP. Specifically, by conducting the Principal Component Analysis (PCA) algorithm on the latent representations of employees, we can

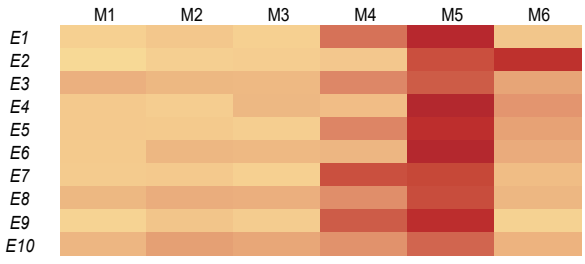


Fig. 9: The attention weights of each month in NNDSP.

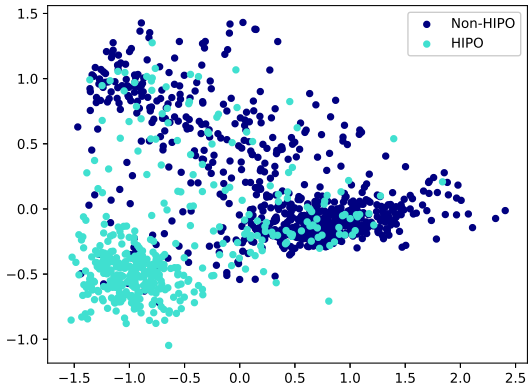


Fig. 10: The PCA visualization of the HIPO representations learned by NNDSP.

plot their representations in a two-dimension space, as shown in Figure 10. From the results, we can find that the HIPOs and non-HIPOs are generally easy to be distinguished. Specifically, the representations of HIPOs are mainly located in the lower left corner, while those of non-HIPOs are located in the top right corner. This result clearly validates the effectiveness of the latent HIPO representations learned by NNDSP.

V. Related Work

The related studies can be grouped into two categories, i.e., HIPO identification and graph-based neural networks.

A. HIPO Identification

How to prospectively identify and develop the HIPOs is always a major concern in the strategic human resource management. Most of existing HIPO identification approaches are still based the subjective experiences of human resource experts, and focus on evaluating the correlations between HIPOs and a variety of talent factors [5], [25]–[27], such as the competencies of communication, teamwork, self-learning and strategic thinking. In addition employee's past performance appraisal is also one of the major data sources which are used in the HIPO identification process [28] [25] [29]. However, these factors are usually manually-selected and suffer from personalized bias due to the subjectivity. Then [5] present analytics algorithms for identifying HIPO employees using organizational databases, to supplement the manually identified ones, which utilize a new method for building a classifier ensemble. Besides, [27] and [30] also explored employees' psychological reactions for HIPO identification. Recently, with the rapid development of management information systems, some data-driven methods are proposed for talent evaluation in some specific scenarios of human resource management, such as the career development prediction [31] [32], person-job fit [33] [34] [35], talent recruitment [36] [37], job skills ranking [38] and turnover prediction [39]. Inspired by these works, in this paper, we propose a neural network approach for quantitatively identifying HIPOs from the newly-enrolled employees, through modeling the dynamics of their organizational social networks, which differs from previous HIPO identification works in terms of both data and method.

B. Graph based Neural Networks

The first work of applying neural network on Euclidean graph data can be dated back to [40], which proposed a method of embedding graphs into the Euclidean space through Recurrent Neural Network (RNN) based supervised learning. Then, researchers have modified the model by using gated recurrent unites and new optimization techniques, which can be further extended to sequential outputs [41]. Recently, with the development of graph signal processing, we have more mathematical tools to make convolution operations available on the non-Euclidean graphs. Thus, the Graph Convolutional Network (GCN) [42] has become one of the most popular techniques for model graph data, which adopts the graph Fourier transforms to apply convolutional operation in the spectral domain. Furthermore, some works have been proposed to reduce the computational complexity of graph convolution manipulation. For example, researchers propose to utilize Chebyshev polynomials and its approximate evaluation scheme to achieve localized filtering [43], and design a first-order approximation of localized filters for semi-supervised classification task [44]. Due to the advanced performance of GCN based techniques for capturing the spatial information in graph data, they have been widely used in applications like traffic network analysis [45] [46], semantic-role labeling [47], event detection, and neural machine translation [48]. In this paper, we propose to utilize GCN to extract the local information of employees in their organizational social network, as a part of their social profiles for HIPO identification.

VI. Conclusion

In this paper, we developed a neural network based quantitative approach for identifying the High Potential talents (HIPOs) from the newly-enrolled employees by modeling the dynamics of their behaviors in organizational social networks. A basic assumption of our approach is that HIPOs usually perform more actively and have higher competencies than their peers to accumulate their social capitals during their work practice. Specifically, we first proposed to model the social profiles of employees with both Graph Convolutional Network (GCN) and social centrality analysis in a comprehensive way. Then, an adaptive Long Short Term Memory (LSTM) network with global attention mechanism was designed to capture the profile dynamics of employees in the organizational social networks during their early career. Finally, extensive experiments performed on real-world data clearly validate the effectiveness of our approach for identifying HIPOs in their earlier career stage. Also, we showed the interpretability of our results.

Acknowledgment

This research was partially supported by grants from the National Natural Science Foundation of China (Grant No. 91746301, 61703386, 61773361).

REFERENCES

[1] R. E. Lewis and R. J. Heckman, "Talent management: A critical review," *Human resource management review*, vol. 16, no. 2, pp. 139–154, 2006.

[2] R. Silzer and A. H. Church, "Identifying and assessing high-potential talent," *Strategy-driven talent management: A leadership imperative*, vol. 28, pp. 213–280, 2010.

[3] N. Dries, T. Vantilborgh, and R. Pepermans, "The role of learning agility and career variety in the identification and development of high potential employees," *Personnel Review*, vol. 41, no. 3, pp. 340–358, 2012.

[4] D. A. Ready, J. A. Conger, and L. A. Hill, "Are you a high potential," *Harvard Business Review*, vol. 88, no. 6, pp. 78–84, 2010.

[5] G. K. Palshikar, K. Sahu, and R. Srivastava, "Ensembles of interesting subgroups for discovering high potential employees," in *PAKDD*, 2016.

[6] C. Beckett and H. Taylor, *Human growth and development*. SAGE Publications Limited, 2019.

[7] G. R. Gonzalez, D. P. Claro, and R. W. Palmatier, "Synergistic effects of relationship managers' social networks on sales performance," *Journal of Marketing*, vol. 78, no. 1, pp. 76–94, 2014.

[8] R. T. Sparrowe, R. C. Liden, S. J. Wayne, and M. L. Kraimer, "Social networks and the performance of individuals and groups," *Academy of management journal*, vol. 44, no. 2, pp. 316–325, 2001.

[9] S. S. Durmuşoğlu, "Merits of task advice during new product development: Network centrality antecedents and new product outcomes of knowledge richness and knowledge quality," *Journal of Product Innovation Management*, vol. 30, no. 3, pp. 487–499, 2013.

[10] E. Whelan, "It's who you know not what you know: a social network analysis approach to talent management," *European Journal of International Management*, vol. 5, no. 5, pp. 484–500, 2011.

[11] R. Yu, Y. Zhang, Y. Ye, L. Wu, C. Wang, Q. Liu, and E. Chen, "Multiple pairwise ranking with implicit feedback," in *CIKM*, 2018.

[12] H. Huang, Y. Dong, J. Tang, H. Yang, N. V. Chawla, and X. Fu, "Will triadic closure strengthen ties in social networks?" *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 12, no. 3, p. 30, 2018.

[13] G. Farnadi, J. Tang, M. De Cock, and M.-F. Moens, "User profiling through deep multimodal fusion," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, 2018, pp. 171–179.

[14] J. M. Kleinberg, "Hubs, authorities, and communities," *ACM computing surveys (CSUR)*, vol. 31, no. 4es, p. 5, 1999.

[15] R. S. Burt, *Structural holes: The social structure of competition*. Harvard university press, 2009.

[16] A. Bavelas, "Communication patterns in task-oriented groups," *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 725–730, 1950.

[17] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.

[18] C. F. Negre, U. N. Morzan, and etc., "Eigenvector centrality for characterization of protein allosteric pathways," *Proceedings of the National Academy of Sciences*, vol. 115, no. 52, pp. E12 201–E12 208, 2018.

[19] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature*, vol. 393, no. 6684, p. 440, 1998.

[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[21] D. Bahdanau, K. Cho, and etc., "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[22] L. Mason, J. Baxter, P. L. Bartlett, and M. R. Frean, "Boosting algorithms as gradient descent," in *Advances in neural information processing systems*, 2000, pp. 512–518.

[23] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[24] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 701–710.

[25] R. Pepermans, D. Vloeberghs, and B. Perkisas, "High potential identification policies: an empirical study among belgian companies," *Journal of Management Development*, vol. 22, no. 8, pp. 660–678, 2003.

[26] N. Dries and R. Pepermans, "Using emotional intelligence to identify high potential: a metacompetency perspective," *Leadership & Organization Development Journal*, vol. 28, no. 8, pp. 749–770, 2007.

[27] J. Gelens, J. Hofmans, N. Dries, and R. Pepermans, "Talent management and organisational justice: Employee reactions to high potential identification," *Human Resource Management Journal*, vol. 24, no. 2, pp. 159–175, 2014.

[28] C. M. Bueno and S. L. Tubbs, "Identifying global leadership competencies: An exploratory study," 2004.

[29] I. Kotlyar, L. Karakowsky, M. Jo Ducharme, and J. A. Boekhorst, "Do "rising stars" avoid risk?: status-based labels and decision making," *Leadership & Organization Development Journal*, vol. 35, no. 2, pp. 121–136, 2014.

[30] R. Slan-Jerusalim and P. A. Hausdorf, "Managers' justice perceptions of high potential identification practices," *Journal of Management Development*, vol. 26, no. 10, pp. 933–950, 2007.

[31] H. Li, Y. Ge, H. Zhu, H. Xiong, and H. Zhao, "Prospecting the career development of talents: A survival analysis perspective," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 917–925.

[32] Q. Meng, H. ping Zhu, K. Xiao, L. Zhang, and H. Xiong, "A hierarchical career-path-aware neural network for job mobility prediction," in *KDD*, 2019.

[33] C. Qin, H. Zhu, T. Xu, C. Zhu, L. Jiang, E. Chen, and H. Xiong, "Enhancing person-job fit for talent recruitment: An ability-aware neural network approach," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2018, pp. 25–34.

[34] D. Shen, H. ping Zhu, C. Zhu, T. Xu, C. Ma, and H. Xiong, "A joint learning approach to intelligent job interview assessment," in *IJCAI*, 2018.

[35] Y. Z. Sun, F. Zhuang, H. ping Zhu, X. Song, Q. He, and H. Xiong, "The impact of person-organization fit on talent management: A structure-aware convolutional neural network approach," in *KDD*, 2019.

[36] Q. Meng, H. ping Zhu, K. Xiao, and H. Xiong, "Intelligent salary benchmarking for talent recruitment: A holistic matrix factorization approach," *2018 IEEE International Conference on Data Mining (ICDM)*, pp. 337–346, 2018.

[37] C. Zhu, H. ping Zhu, H. Xiong, P. Ding, and F. Xie, "Recruitment market trend analysis with sequential latent variable models," in *KDD*, 2016.

[38] T. Xu, H. Zhu, C. Zhu, P. Li, and H. Xiong, "Measuring the popularity of job skills in recruitment market: A multi-criteria approach," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[39] M. Teng, H. ping Zhu, C. Liu, C. Zhu, and H. K. Xiong, "Exploiting the contagious effect for employee turnover prediction," in *AAAI*, 2019.

[40] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *2005 IEEE International Joint Conference on Neural Networks*, vol. 2. IEEE, 2005, pp. 729–734.

[41] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.

[42] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," *CoRR*, vol. abs/1312.6203, 2013.

[43] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *NIPS*, 2016.

[44] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *CoRR*, vol. abs/1609.02907, 2016.

[45] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *ICLR*, 2017.

[46] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "High-order graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," *arXiv preprint arXiv:1802.07007*, 2018.

[47] D. Marcheggiani and I. Titov, "Encoding sentences with graph convolutional networks for semantic role labeling," in *EMNLP*, 2017.

[48] J. Bastings, I. Titov, W. Aziz, D. Marcheggiani, and K. Sima'an, "Graph convolutional encoders for syntax-aware neural machine translation," in *EMNLP*, 2017.