Tracking the Evolution of Member Roles in Practice Communities with Integrated Knowledge Graphs and Mining the Path to Building the Identity of Entrepreneurial Mentors

Xiuqin Zhang
School of Innovation and Entrepreneurship, Nanjing University of Industry Technology, Nanjing 210023, Jiangsu, China
Corresponding author: Xiuqin Zhang (e-mail: yuyongshenf@163.com)

ABSTRACT Member-role evolution and entrepreneurial mentor identification in Practice Communities have the following problems: high difficulty in dynamic tracking, insufficient integration of multi-dimensional features and implicit evolutionary path. Therefore, this paper proposes a two-layer analysis framework based on Knowledge Graph (KG): Firstly, a multi-dimensional knowledge graph based on member interaction, knowledge contribution and social network is constructed and member-role-behavior triplets are defined by entity extraction and relation mining; Secondly, a kind of Temporal Graph Convolutional Network (TGCN) is designed to capture trajectory of role evolution and PageRank algorithm is used to quantify node influence and discover the evolutionary path from peripheral participant to core mentor; Finally, a path mining algorithm is used to extract key transformation node and behavioral feature sequence. Experiments are based on three years of data from an Practice Communities, constructing a knowledge graph containing member nodes and 15293 relation edges. The average node degree is 10.74, the network density is 0.0038, and the clustering coefficient reaches 0.52, indicating that the community exhibits obvious smallworld characteristics. The results showed that the role classification module divided members into five roles: latent observers, casual participants, active contributors, domain experts, and entrepreneurial mentors, with a classification accuracy of 87.3%. Evolutionary trajectory analysis tracked the complete growth process of 217 members who eventually became entrepreneurial mentors, experiencing an average of 4.2 role transitions.

INDEX TERMS knowledge graph, practice community, role evolution, temporal graph convolution, entrepreneurial mentor

I. INTRODUCTION

BEING regarded as a significant organizational structure of knowledge sharing and collaborative learning, interaction and knowledge flow between the Practice Communities members influence dynamic development of individual roles. The identity construction process of entrepreneurial mentors is a multi-dimensional transformation, i.e., knowledge accumulation, expansion of social relationships, and enhancement of influence are the key roles of the community. Current literature is largely in a static frame of reference to examine the nature of member roles, and has no dynamic tracking system to trace the evolutionary paths. The trajectories through which members transition from peripheral participants to core mentors remain hidden within complex interaction networks. The conventional approaches cannot adequately represent the most important nodes and driving forces of role transformation, which limits the effective development of mentor resources in the entrepreneurial communities [1,2].

Knowledge graph technology offers a formal representation approach to describe the multidimensional connections of a group of practice, and is able to incorporate heterogeneous information, including member qualities, behavior patterns, and network structure. Graph neural networks have demonstrated excellent representation learning on graph-structured data and the temporal modeling sub-module is able to model how the state of nodes varies with time. A combination of knowledge graphs and temporal graph convolutional networks can provide dynamic tracking and pattern recognition of role evolution in a single framework [3,4]. Path mining algorithms can identify overall identity constructions patterns by following the historical course of effective cases, offering a course of action guidance path to mentor training.

The paper proposes that a method based on knowledge graphs can be applied to trace the process of evolution of roles of the members in a Practice Communities and to mine the trajectory of identity-construction process of an entrepreneurial mentor. It is a multi-dimensional graph of knowledge including not only interaction among the members but also contribution of knowledge and social networks. Temporal graph convolutional network model is trained to learn dynamic trajectory of role transformation and then a path mining algorithm is employed to mining important turning points and sequences of behavioral features. Three years data of a startup community are used to perform experiments to validate the efficacy of our method and then five kinds of common process evolution are identified and the internal process of identity construction of an entrepreneurial mentor [5,6] is mined out. The quantitative decision support that the results of our research can provide are that the mentor can precisely cultivate the mentor resources in startup communities and promote the intelligent construction of knowledge management in communities of practice.

II. RELATED WORK

The development of roles of the members of a Practice Communities is a research area that is interested in the process and mechanism through which individuals change peripheral members to core members. Initial research developed a valid marginal participation model that is grounded on the social learning theory, which outlines how members slowly come to be a part of the community through observation, imitation and practice. Regarding the classification of roles, scholars have suggested different classification systems, including hierarchical systems like latent participants, novices, contributors and experts [7,8]. Evolutionary driving factors are analyzed in the following dimensions: the ability to acquire knowledge, social capital accumulation, the level of motivation to participate. Current approaches largely depend on questionnaire surveys and interview data, which are not capable of extracting large scale behavioral data, thus may not be able to uncover the implicit evolutionary patterns.

The application of knowledge graphs to the analysis of social networks is a radical way to define the tricky relationships. Entity recognition, relation extraction and attribute completion are ways; graph representation learning ways project discrete structures into continuous space. GNNs get the neighbor information via message passing architectures, and the different types of (e.g., graph convolutional networks and graph attention networks) have achieved great performance on node classification and link prediction tasks. Temporal graph neural networks add a time dimension to the evolution process of dynamic processes in models using recurrent units or time attention mechanism to learn the change of node states. They could be used in the field of social network analysis, recommender systems, bioinformatics, etc. while few research works have been done on how the roles in communities of practice change over time.

The studies of identity construction of entrepreneurial mentors concentrate on the mechanism of formation of the mentor role and the factors that affect its formation. The theoretical views are social identity theory, role theory and professional knowledge development frameworks. The relationship between mentors knowledge structure, mentoring behaviour, and network position in identity establishment have been investigated empirically [9,10]. The measurement techniques include self-report measures, peer rating, and network centrality measures. Existing literature mostly focuses on identity construction under formal mentoring systems, lacking systematic analysis of the natural emergence process of mentors in open entrepreneurial communities. Path identification can be primarily qualitative induction with lack of quantitative tracking tools.

III. METHODS

A. METHODS FOR CONSTRUCTING MULTIDIMENSIONAL KNOWLEDGE GRAPHS

The multidimensional knowledge graph is built based on the interactive logs of an Practice Communities as the source of data, including the records of the activities of the members of the network, including posting, commenting, and sharing resources via web crawling. The entity extraction module takes a BERT pre-trained model to detect three types of entities: member identity, knowledge domain, and activity type. It uses BiLSTM-CRF sequence labeling structure to identify entity boundaries and removes low-frequency noise entities with the TF-IDF algorithm. Interactive edges are created at the relationship construction level using the reply relationships among the members, collaborative edges are created using the jointly participated projects, and propagation edges are created using the knowledge citation records. Every edge has three-dimensional weights of attributes, which are: timestamp, interaction frequency, and content similarity. The graph combines both social network topological and semantic information, with TransE learning entity embedding representations and computes potential associations based on cosine similarity to fill in sparse connections. To ensure a dynamic update mechanism, the construction process is performed by adding new interactive data periodically (once a week) that are generated, and the node and edge attribute data are stored in the Neo4j graph database. To identify the entrepreneurial mentors, the special characteristics of knowledge contribution, the number of citations, and the percentage of answers are specially denoted, and the multi-layer heterogeneous graph is built that also incorporates the evolutionary history of members.

B. ROLE EVOLUTION TRACKING MODEL BASED ON TEMPORAL GRAPH CONVOLUTIONAL NETWORKS

Given a trained multidimensional knowledge graph as input, our temporal graph convolutional network model segments the timeline into sequences of monthly snapshots. The architecture adopts a serial connection where the spatial features are extracted first; specifically, the GCN layers perform spectral domain convolution on each snapshot to generate high-dimensional node embeddings, which are then fed as a time-series sequence into the GRU layers to capture temporal dependencies. The TGCN model processes temporal snapshots. For each snapshot, the spectral graph convolution is defined as:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{1}$$

where $A = \tilde{A} + I_n$ represents the adjacency matrix with self-loops, and $\tilde{D}$ is the corresponding degree matrix. We employ the symmetric normalized Laplacian matrix to ensure stable information propagation. This first-order approximation of Chebyshev polynomials allows for efficient capture of the multi-hop neighborhood information across three layers.

The temporal modeling module consists of Gated Recurrent Units (GRU). GRU was selected over LSTM because it features a more streamlined structure with fewer parameters, reducing the risk of overfitting on the sparse interaction data typical of CoPs while maintaining comparable performance in capturing long-term dependencies. Preliminary baseline experiments indicated that TGCN-GRU achieved a 2.4% higher F1-score compared to TGCN-LSTM and demonstrated better convergence speed than Transformer-based architectures in this specific small-sample temporal graph context. The contribution scores of different time periods are output by the attention mechanism that assigns higher contribution scores to critical time points of role transitions of mentors.

The role classifier outputs the member into five different role categories: latent observer, casual participant, active contributor, domain expert, entrepreneurial mentor. The training of the model uses a cross-entropy loss that needs to be minimized to increase the accuracy of classification, and adds to it a role transition smoothing constraint that avoids large drops in the label.

C. ENTREPRENEURIAL MENTOR IDENTITY CONSTRUCTION PATH MINING ALGORITHM

Path mining algorithm begins with member nodes already identified as entrepreneurial mentors, and follows them through their entire evolutionary history, building a role sequence over time. The algorithm is based on depth-first search over the time edges of the knowledge graph, which logs the role transition nodes that each member went through when joining the community, to receiving mentor status. The turning point identification module finds the rate of change of features between adjacent time windows, which indicates significant events when the rate of increase in knowledge contribution is greater than a threshold or when the network centrality has suddenly changed.

The behavioral feature extractor compares the patterns of interaction prior to and after turning points, statistically comparing quantitative measures of frequency of posting, depth of replies and frequency of cross domain collaboration. Pattern mining is common and FP-Growth algorithm is applied to identify high

frequency combinations of behavior and minimum support is set to 20 and typical patterns are identified including continuous knowledge output and cross-domain collaboration. Cluster analysis is a merging of similar evolutionary processes, in which sequence similarity is calculated by the dynamic time warping algorithm. K-Means grouping classifies the paths into five types; linear progression, spiral growth, platform innovation, multi-role change, and the fast leap. Path validation uses leave-one-out cross-validation, and individual paths are extracted to evaluate the accuracy of model prediction, and the average bias in prediction is controlled to 0.16 months. The mining outputs are represented as a directed acyclic graph (DAG) with a relationship of path branching and convergence.

D. MODEL VALIDATION METHODS

The model evaluation system builds an index set based on three dimensions: role classification accuracy, evolutionary tracking stability and path mining completeness. Role recognition evaluation employs precision, recall, and F1 score to evaluate the classification performance of five role types, and presents both macroaveraging and microaveraging aggregation to deal with class imbalance. To determine the quality of evolutionary tracking, a temporal consistency index is used, which compares the entropy of the transition probability distribution of prediction outputs over a continuous time interval; the lower the entropy, the more steady is the trajectory prediction. Path mining analysis includes a coverage index that uses statistical analysis of the percentage of significant turning points identified by the algorithm to the events labelled manually at the same time the path similarity scores are computed to confirm the sanity of the clustering.

The validation procedure uses stratified sampling to divide the data based on their time of joining the group and after that, it is divided into training, validation and test samples in a ratio of 7:2:1. Cross-validation employs a 5 fold time series partitioning to guarantee the continuity of each fold in time and prevent information leakage. The control experiment establishes fixed graph neural networks, recurrent neural networks and random walk algorithms as baseline algorithms. Table 1 presents the definitions and calculation techniques of the various methods in the different dimensions of evaluation with five columns, the name of the indicators, mathematical formula, value range, optimal direction and scenario to use.

TABLE 1. Model Evaluation Index System

Indicator Name	Calculation method	Range of values	Optimal direction	Applicable Scenarios
Character Classification F1 Value	Harmonic average of precision and recall	0-1	The higher the better	Character recognition accuracy evaluation
Temporal consistency index	The negative value of the entropy of the transition probability distribution	negative infinity to 0	The closer to 0, the better.	Evolutionary trajectory stability verification
Inflection point coverage	Number of identified events divided by number of labeled events	0-1	The higher the better	Completeness of key node mining
Path similarity score	Dynamic time warped distance reciprocal	0-1	The higher the better	Clustering quality assessment
Prediction time deviation	Difference between predicted and actual time	0 to positive infinity	The lower the better	Evolution time prediction accuracy

IV. RESULTS AND DISCUSSION

A. KNOWLEDGE GRAPH CONSTRUCTION RESULTS AND STATISTICAL ANALYSIS

Complete interaction records of a startup community were collected between January 2023 and December 2025, with heterogeneous data including information on member registration, the content of posts, replies on comments, sharing of resources, and collaboration in projects. The entity extraction module discovered 2,847 valid member entities out of 81,246 text records. The best-statistical features of the knowledge graph are represented in Table 2. The mean degree of nodes is 10.74, the network density is 0.0038 and the clustering coefficient is 0.52 which shows that the community has evident small-world properties. The analysis based on time dimensions indicates that the graph size is increasing at a rate that is increasing at an accelerated rate.

TABLE 2. Statistical Characteristics of Knowledge Graph Topology

Statistical indicators	Numerical values	Statistical indicators	Numerical values
Total number of nodes	2847	Average node degree	10.74
Total number of relation edges	15293	Maximum node degree	237
Network density	0.0038	Clustering coefficient	0.52
Average path length	4.26	Power-law distribution index	2.31
Number of connected components	12	Percentage of maximum connected components	94.3%

Table 2 shows that the degree distribution follows a power-law distribution, and the average path length of 4.26, combined with a high clustering coefficient of 0.52, verifies the small-world network characteristics, indicating high information propagation efficiency among members. The maximum node degree of 237 belongs to the official community manager account, which serves as a central hub for information dissemination. To ensure data validity, we conducted anomaly detection using the Forest-based out-of-distribution method, confirming that while this hub node significantly impacts connectivity, its interaction patterns (replies and resource sharing) are consistent with human-led community facilitation rather than automated bot behavior. The fact that the largest connected component covers 94.3% of the nodes indicates good overall connectivity within the community, with only 12 isolated small groups existing.

B. ROLE EVOLUTION TRAJECTORY RECOGNITION AND TYPICAL PATH CLASSIFICATION

A time-series graph convolutional network model was used as input for 36 months of snapshot data to identify the role labels of 2847 members in different time windows. After 200 training cycles, the model was converged, and the loss in the validation set was 0.143. The role classification module has classified the members into five roles: latent observers, casual participants, active contributors, domain experts and entrepreneurial mentors with 87.3% classification accuracy. The entire growth process of 217 individuals who eventually became entrepreneurial mentors was monitored through evolutionary trajectory analysis, and had on average 4.2 role transitions. Table 3 presents the statistical characteristics of the five typical evolutionary paths, with the linear promotion path accounting for the highest proportion at 38.7%.

TABLE 3. Statistics on the Evolutionary Path Characteristics of Five Typical Roles

Path type	Number of members	Percentage	Average evolutionary cycle (months)	Average number of character stages
Straight-line promotion	84	38.7%	19.6	3.8
Spiral Ascent	56	25.8%	24.3	5.1
Platform Breakthrough	41	18.9%	27.8	4.6
Multi-role switching type	23	10.6%	31.2	6.3
Rapid transition type	13	6.0%	12.4	2.7

Table 3 shows that the linear advancement path has the highest proportion and the shortest evolutionary cycle, with members steadily advancing in role levels in a fixed order. The spiral ascent path involves 5.1 role stages, with brief regressions but an overall upward trend. The rapid leap path completes mentorship in just 2.7 role stages, compressing the evolutionary cycle to 12.4 months, but has the smallest sample size at only 6.0%. The multi-role switching path has the longest evolutionary cycle at 31.2 months, with frequent transitions between different roles resulting in the highest path complexity.

C. ANALYSIS OF KEY TURNING POINTS IN THE CONSTRUCTION OF THE ENTREPRENEURIAL MENTOR IDENTITY

Path mining algorithm followed stepwise the entire evolutionary history of 217 entrepreneurial mentors and found three major types of inflection points. The first knowledge sharing inflection point is when the member starts to give back and cease to only receive knowledge; once this has occurred, the out-degree of the member in the knowledge graph grows exponentially, and the diversity of neighboring nodes is enhanced. The cross-domain collaboration inflection point is defined by the breaking of the initial professional boundaries, the involvement of members in several thematic projects to build the cross-domain links, and the increase in the level of centrality. The maintained interaction maintenance inflection point indicates that the members have created a stable pattern of activity with a lower fluctuation coefficient of the frequency of interactions and shorter response time. The temporal pattern of inflection points has a definite regularity: the initial knowledge sharing is focused between the 5th and 8th month of the community membership, cross-domain collaboration is focused between the 12th and 17th month, and the pattern of the continuation of interaction is established after the 20th month. Patterns of combination of inflection points vary according to the type of evolutionary path; the linear progression type is strictly in order of the three inflection points, and the type of rapid leap type only has the inflection points of knowledge sharing and cross-domain cooperation.

V. CONCLUSION

The present paper forms a two-layer knowledge graph analysis system that is useful in overcoming the difficulties associated with tracing the development of member roles within the entrepreneurial communities and excavating the trail towards the building of the entrepreneurial mentor identities. The multi-dimensional knowledge graph combines member interaction, knowledge contribution and social network properties; the temporal graph convolutional network reflects the dynamic evolutionary patterns of role transformation and the path mining algorithm discloses the inherent mechanism of transformation of peripheral participants into core mentors. Experimental findings demonstrate that this approach is capable of precisely determining important turning points and sequence of behavioral features, which can offer a measurable decision-making groundwork to the accurate development of mentors in entrepreneurial groups. The weakness of the approach is that it analyzes only the data on one of the entrepreneurial communities, and its transferability across different platforms should be additionally proven. Future efforts will be extended to collaborative study of multiple communities, provide external sources of knowledge to enhance the dimensions of features, experiment with techniques of transfer learning to enhance the performance of models of generalization, and combine techniques of causal inference to analyze the driving forces of role evolution in depth.

AUTHOR CONTRIBUTIONS

Xiuqin Zhang designed, collected and analyzed the data, and drafted the manuscript. Xiuqin Zhang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Xiuqin Zhang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research was supported by, 2024 Special Research Project on the Digital Transformation of Higher Education and the Practice of Educational Modernization: «Research on the Evaluation of Implementation Effects and Countermeasures of Innovation and Entrepreneurship Online Education in Vocational Colleges».

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] B. Hutchins and R. Boyle, "A Community of Practice," Changing Sports Journalism Practice in the Age of Digital Media, pp. 4-20, 2020, doi: 10.4324/9780429344886-2.

[2] The Community of Practice, "Advances in Educational Technologies and Instructional Design," 2022:181-201, doi: 10.4018/978-1-7998-7130-9.ch008.

[3] K. Smyth, "The LD4D community of practice," Bulletin de lOIE, 2021, doi: 10.20506/bull.2021.1.3265.

[4] Redactie, "IPOS als Community of Practice," Onderwijs En Gezondheidszorg, vol. 44, pp. 11-11, 2020, doi: 10.24078/oeng.2020.10.126033.

[5] J. Mcdonald and M. Mercieca B, "What Is a Community of Practice? 2021,", doi: 10.1007/978-981-33-6354-0_1.

[6] L. Sun, Z. Tao, Y. Li, et al., "ODA: Observation-Driven Agent for integrating LLMs and Knowledge Graphs," 2024, doi: 10.18653/v1/2024.findings-acl.442.

[7] R. Slavin, R. Williams, and C. Zimmermann, "Building a Community of Practice," Activating the Art Museum, pp. 167-174, 2023, doi: 10.5771/9781538158555-167.

[8] A. Kraemer and D. Steider, "Building a Community of Practice," Sharing Less Commonly Taught Languages in Higher Education, pp. 233-246, 2023, doi: 10.4324/9781003349631-22.

[9] C. Yu, "Narrative community in a community of practice," Online Collaborative Translation in China and Beyond, pp. 61-84, 2022, doi: 10.4324/9781003024200-4.

[10] E. Mccoll and N. Wilson, "Community of practice and Royal Assent," Dental Update, vol. 51, no. 7, pp. 455-456, 2024, doi: 10.12968/denu.2024.51.7.455.