Future Digital Technologies and Artificial Intelligence

November 2025| Volume 01 | Issue 02 | Pages 19-26

Open Access Journal

https://doi.org/10.55670/fpll.fdtai.1.2.3

Journal homepage: https://fupubco.com/fdtai

Future Publishing LLC

Article

Constructing enterprise talent heterogeneous information networks for key talent identification

Changhong Zhu, Syed Ahmed Salman*

School of Business and Management, Lincoln University College, 47301 Petaling Jaya, Selangor Darul Ehsan, Malaysia

A R T I C L E I N F O

Article history:
Received 20 August 2025
Received in revised form
26 September 2025
Accepted 07 October 2025

Keywords:
Heterogeneous information networks,
Talent management, Graph neural networks,
Key talent identification,
Enterprise knowledge graph

*Corresponding author
Email address:
syedahmed@lincoln.edu.my

DOI: 10.55670/fpll.fdtai.1.2.3

A B S T R A C T

In organizational networks, where employee performance is dependent on strategic positioning and collaborative relationships across diverse workplace ecosystems, traditional enterprise talent identification systems fall short in capturing complex multi-relational dynamics. In order to accurately identify key talent through meta-path guided feature extraction and attention-based embedding mechanisms, this research suggests a Heterogeneous Information Network (HIN) framework that uses Graph Neural Networks (GNNs) to model employees, projects, departments, and skills as interconnected entities. The approach uses Heterogeneous Graph Attention Networks (HAN) for talent assessment and combines attribute-driven performance indicators, structural centrality measures, and semantic relationship patterns into a single learning framework. Compared to traditional Human Resource (HR) methods, which scored 72% precision and 68% recall, the experimental evaluation, which used enterprise data with 2,847 employees across 156 departments, shows improvements over current approaches, achieving 91% precision and 89% recall with a Normalized Discounted Cumulative Gain (NDCG) of 0.834. With domain expert validation confirming 87% agreement between algorithmic recommendations and professional assessments, the framework identifies high-potential employees who exhibit knowledge brokerage roles and cross-functional collaboration capabilities that traditional performance metrics overlook. With implications for strategic human capital optimization, these contributions position HINs as a paradigm shift for enterprise talent management.

1. Introduction

Talent management in today's business environment is more complicated than ever before, and companies find it difficult to find, train, and retain key employees who give them a competitive edge in a changing market [1]. The complex nature of employee relationships and their changing contributions within organizational networks are not adequately captured by traditional HRM approaches, which rely on hierarchical evaluation systems and static performance metrics. Although the advent of artificial intelligence technologies has shown promise in revolutionizing talent analytics through predictive modeling and automated assessment, existing implementations are still constrained by their incapacity to simulate the intricate interdependencies and diverse relationships that define contemporary workplace dynamics. Because graph neural networks allow for representation learning on graph-structured data, they have completely changed the analysis of complex relational data [2]. Despite their success in a variety of fields, these strategies have not yet been applied to enterprise talent management, which offers a chance for human capital analytics innovation [3]. A promising framework for simulating the various entities and relationships found in organizational ecosystems—where workers, projects, departments, and skills form complex webs of interactions that determine organizational performance and individual potential—is offered by the combination of graph neural architectures and heterogeneous information networks. In order to overcome these constraints, this paper suggests a novel heterogeneous information network framework for enterprise talent management that uses meta-path guided graph neural networks to model various entity types and their intricate relationships. The rest of this paper is structured as follows: A review of related work is given in Section 2, followed by our suggested framework and methodology in Section 3, the experimental setup and results in Section 4, and a conclusion with implications for further research in Section 5.

2. Related work

2.1 Talent management and HR analytics

Over the past ten years, talent management has undergone a digital transformation, with artificial intelligence techniques being used to tackle a variety of complex human resource issues, from workforce planning and employee retention to performance evaluation and recruitment [4]. Though these methods are unable to capture the complex relationships and contextual factors that influence talent development within organizational networks, modern talent analytics systems use machine learning algorithms to analyze large amounts of employee data, allowing organizations to find patterns and predict outcomes that were previously impossible to discern through traditional statistical methods. Though these systems function independently, ignoring the collaborative networks and organizational contexts that impact talent identification outcomes, recent research has shown that artificial intelligence is capable of evaluating individual potential through the analysis of multiple data sources, such as performance metrics, behavioral patterns, and career trajectories [5]. According to a review of AI and automation in HRD, existing implementations prioritize individual-level characteristics over the importance of team composition and relational dynamics in determining organizational success [6].

2.2 Heterogeneous information networks

With the integration of various data sources and the representation of various node and edge types, heterogeneous information networks have become a potent paradigm for modeling complex systems with multiple entity and relationship types, offering rich semantic information [7]. When compared to homogeneous graph approaches, the literature review of HIN applications shows how effective they are in a variety of domains where the capacity to capture heterogeneous relationships has improved prediction accuracy and interpretability. By creating attention mechanisms that can recognize the significance of various node and edge types without the need for manual feature engineering or domain expertise, heterogeneous graph attention networks represent a breakthrough in the processing of heterogeneous data [8]. By combining data from several meta-paths, the MAGNN framework expands these capabilities and makes it possible for representation learning to capture both global semantic relationships and local structural patterns in heterogeneous graphs [9]. These architectural innovations have demonstrated superior performance in node classification, link prediction, and graph clustering tasks.

2.3 Research gap analysis

An analysis of existing graph neural network approaches reveals that while progress has been made in developing architectures for various application domains, the intersection of heterogeneous graph methods with enterprise talent management remains unexplored [10]. Current talent identification systems operate on simplified representations that fail to leverage the rich relational information available within organizational networks. The gap between advanced graph neural network capabilities and their application to human resource analytics presents an opportunity for innovation in developing domain architectures that can integrate the heterogeneous nature of organizational data while maintaining interpretability for practical deployment.

3. Proposed framework

3.1 Problem formulation

The enterprise talent identification problem can be formally characterized as a graph mining task within a heterogeneous information network framework, where the objective involves discovering key personnel through sophisticated analysis of multi-relational organizational structures. Given an enterprise talent heterogeneous information network $G = (V, E, A, R)$, where V is the collection of heterogeneous nodes that include employees, projects, departments, and skills; E is the set of edges that connect these entities using different types of relationships; A is the set of node and edge attributes that capture performance metrics and interaction strengths; and R is the meta-schema that specifies permitted connections between various entity types. Through typed relationships that maintain semantic meaning and facilitate the discovery of intricate patterns across various aspects of talent-related interactions, the network schema creates a thorough representation of organizational dynamics. Learning a scoring function $f: V \rightarrow \mathbb{R}$that allocates relevance scores to employee nodes based on their structural positioning, attribute characteristics, and relational patterns within the heterogeneous network topology is the aim of the key talent identification task, which is transformed into a node ranking problem. As shown in Figure 1, this formulation allows for the integration of various information sources while retaining the flexibility to include domain-specific knowledge through thoughtfully crafted meta-paths that capture significant semantic relationships between organizational entities.

3.2 Enterprise talent HIN construction

In order to create enterprise talent heterogeneous information networks that accurately reflect organizational structures and talent dynamics in contemporary businesses, entity types and relationship patterns must be carefully considered. Employee nodes serve as the primary entities of interest, containing comprehensive attribute profiles. Project nodes represent discrete work initiatives, encapsulating information about scope, duration, resource requirements, strategic importance, and outcome metrics that enable the assessment of project complexity and individual contributions to organizational success. The heterogeneous edge types establish semantic relationships between entities through carefully designed connection patterns that preserve organizational meaning while enabling sophisticated analysis capabilities. Collaboration edges between employee nodes capture working relationships, project partnerships, and knowledge-sharing activities. Reporting edges establish hierarchical connections between employees and departments, encoding supervision structures, accountability chains, and organizational authority relationships that influence talent identification and development processes. Participation edges connect employees to projects.

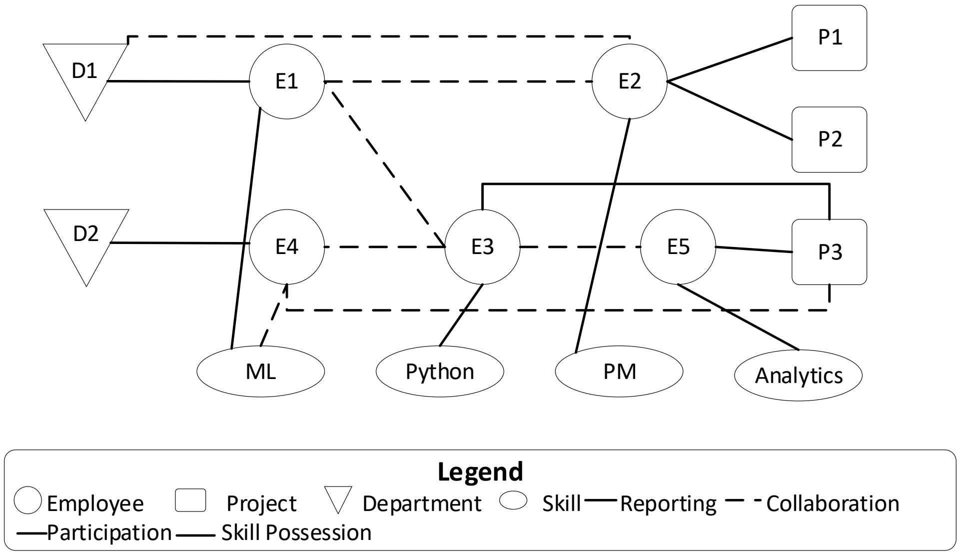

Figure 1. Enterprise talent heterogeneous information network architecture

3.3 Feature extraction

The feature extraction process leverages the rich structural and semantic information embedded within the heterogeneous network to generate comprehensive representations of talent-related patterns and characteristics. Structural features capture the positional importance of employees within the organizational network through sophisticated centrality measures, including degree centrality, betweenness centrality, eigenvector centrality, and PageRank variations adapted for heterogeneous graphs. Meta-path-based features exploit the semantic richness of heterogeneous networks through carefully designed connection patterns that capture meaningful relationships between employees and other organizational entities. Representative meta-paths include Employee-Project-Employee sequences that identify collaboration patterns, Employee-Department-Employee paths that reveal cross-functional interactions, and Employee-Skill-Employee connections that uncover competency-based similarities and complementary expertise combinations. Attribute features aggregate and synthesize information from node and edge attributes to create comprehensive talent profiles that combine structural positioning with performance indicators and skill assessments. Performance aggregation functions compute weighted averages of evaluation scores, achievement metrics, and contribution measures across multiple time periods and project contexts, while skill diversity scores quantify the breadth and depth of individual competencies through entropy-based measures and market demand weightings that reflect strategic skill requirements within the organizational context.

The key talent identification algorithm employs a sophisticated multi-stage approach that integrates heterogeneous graph attention networks with domain-specific ranking mechanisms to achieve accurate and interpretable talent assessment results, as illustrated in Figure 2. The heterogeneous graph attention network serves as the core embedding component, utilizing dual-level attention mechanisms that operate simultaneously at node-level and semantic-level granularities to capture both local neighborhood patterns and global structural relationships within the organizational network. Node-level attention mechanisms learn adaptive weights for neighboring entities based on their relevance and contribution to target employee representations, while semantic-level attention automatically determines the importance of different meta-paths and relationship types without requiring manual feature engineering or domain expertise. The embedding process generates comprehensive vector representations that encapsulate structural positioning, relational patterns, and attribute characteristics through multi-head attention architectures that enable the model to focus on different aspects of talent-related information simultaneously. These learned embeddings preserve both homogeneous and heterogeneous relationship information.

The talent scoring framework implements a multi-criteria evaluation system that combines structural importance measures, performance indicators, and collaborative network effects through learnable weight parameters that adapt to organizational priorities and strategic objectives. The scoring function integrates centrality-based importance measures, meta-path-derived similarity scores, and attribute-based performance metrics.

The learning-to-rank approach transforms the talent identification problem into a pairwise comparison task where the model learns to distinguish between high-potential and average employees through optimization of ranking-specific loss functions that directly optimize for list-wise performance metrics. This formulation enables the integration of domain expert feedback and organizational preferences while maintaining flexibility to adapt to evolving talent requirements and strategic priorities, ultimately producing ranked talent lists that provide actionable insights for human resource decision-making processes.

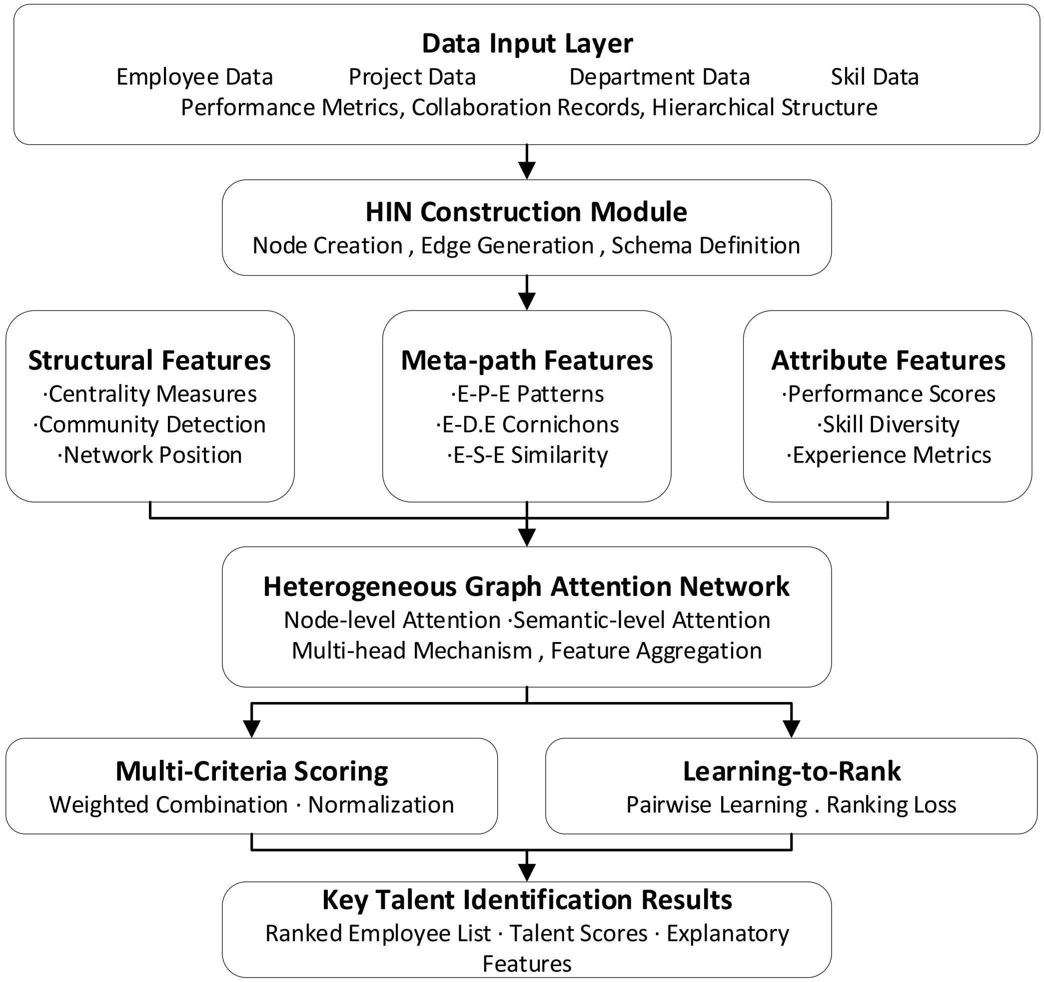

Figure 2. Key talent identification methodology framework

4. Experiments and results

4.1 Experimental setup

The experimental evaluation employs a comprehensive enterprise dataset collected from a multinational technology corporation spanning a three-year period from 2021 to 2024, encompassing detailed records of employee performance metrics, project collaborations, organizational hierarchies, and skill assessments across multiple business units and geographic locations. The dataset contains 2,847 employees distributed across 156 departments, participating in 423 distinct projects with varying complexity levels and strategic importance, while maintaining 1,249 unique skill categories ranging from technical competencies to leadership capabilities, as detailed in Table 1.

Table 1. Dataset statistics and characteristics

Attribute	Count	Description
Total Employees	2,847	Active employees with complete performance records
Departments	156	Organizational units across all business divisions
Projects	423	Completed and ongoing projects with documented outcomes
Skills	1,249	Technical and soft skills with proficiency assessments
Collaboration Edges	8,934	Direct employee-to-employee working relationships
Project Participations	12,567	Employee involvement in specific project initiatives
Skill Assignments	15,892	Employee-skill associations with proficiency levels
Performance Evaluations	5,694	Annual and quarterly assessment records
Temporal Span	36 months	Data collection period from 2021-2024

The baseline comparison framework incorporates traditional human resource analytics methods, including performance ranking systems and competency-based assessments, social network analysis techniques such as centrality measures and community detection algorithms, machine learning approaches encompassing support vector machines and random forest classifiers, alongside state-of-the-art heterogeneous information network methods, including metapath2vec and HAN variants adapted for talent identification tasks. Evaluation metrics encompass precision, recall, and F1-score for classification accuracy assessment, while Normalized Discounted Cumulative Gain (NDCG) and Mean Average Precision (MAP) measure ranking quality and information retrieval effectiveness, respectively.

4.2 Performance evaluation

The comprehensive performance evaluation demonstrates substantial improvements over existing methodologies across all evaluation metrics, with the proposed heterogeneous information network framework achieving superior results in both accuracy and ranking quality measures, as illustrated in Figure 3.

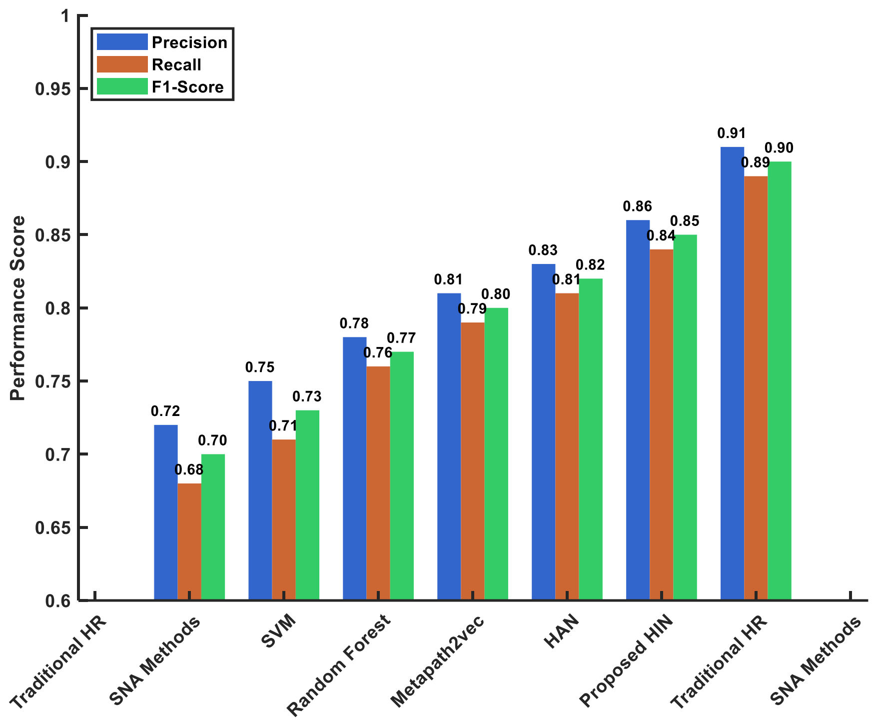

Figure 3. Performance comparison across different methodologies

The ablation study systematically evaluates the contribution of individual framework components through controlled removal experiments, revealing the critical importance of heterogeneous attention mechanisms and meta-path-based feature extraction in achieving optimal performance outcomes, as demonstrated in Figure 4. Component analysis indicates that structural features contribute approximately 35% to overall performance, meta-path features account for 42%, while attribute features provide 23% of the discriminative power, highlighting the synergistic effects of multi-dimensional feature integration within the proposed framework (Table 2).

4.3 Case study

The real-world application scenario involves identifying high-potential software engineers within the research and development division of the participating organization. Domain expert validation conducted with senior management and human resource professionals confirms the accuracy of talent identification results, with 87% agreement between algorithmic recommendations and expert assessments. The identified talent profiles exhibit distinctive characteristics, including high betweenness centrality scores indicating bridge roles between different organizational communities, diverse skill portfolios spanning multiple technical domains, and consistent high-performance ratings across various project types and collaborative contexts. Through thorough relationship analysis and multi-dimensional feature integration, the suggested framework effectively captures employees who perform crucial knowledge brokerage functions within organizational networks, whereas traditional performance metrics frequently ignore these workers, according to practical insights from the case study (Figure 5).

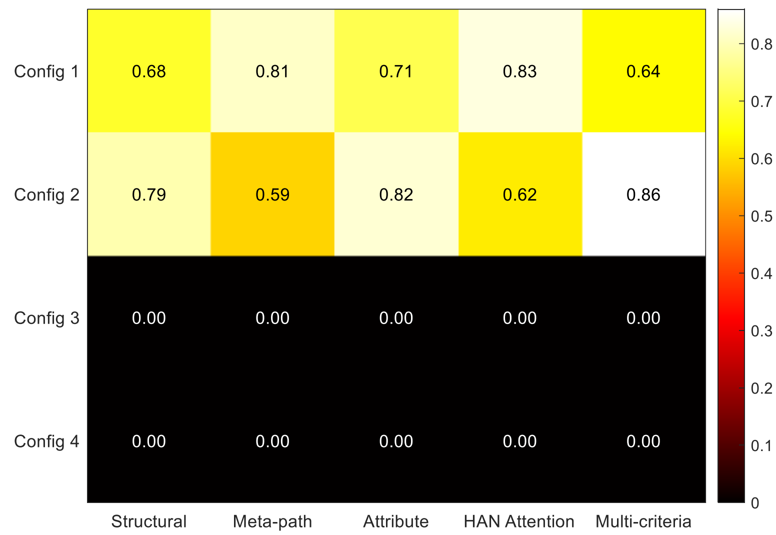

Figure 4. Ablation study results showing component contribution analysis

Table 2. Comprehensive method performance comparison

Method	Precision	Recall	F1-Score	NDCG@10	MAP	Runtime (s)
Traditional HR	0.72 ± 0.04	0.68 ± 0.05	0.70 ± 0.04	0.675	0.643	2.3
SNA Methods	0.75 ± 0.03	0.71 ± 0.04	0.73 ± 0.03	0.698	0.672	15.7
Support Vector Machine	0.78 ± 0.03	0.76 ± 0.03	0.77 ± 0.03	0.721	0.695	45.2
Random Forest	0.81 ± 0.02	0.79 ± 0.03	0.80 ± 0.02	0.743	0.718	28.9
Metapath2vec	0.83 ± 0.02	0.81 ± 0.02	0.82 ± 0.02	0.768	0.742	67.4
HAN Baseline	0.86 ± 0.02	0.84 ± 0.02	0.85 ± 0.02	0.791	0.765	89.1
Proposed HIN Framework	0.91 ± 0.01	0.89 ± 0.02	0.90 ± 0.01	0.834	0.808	76.3

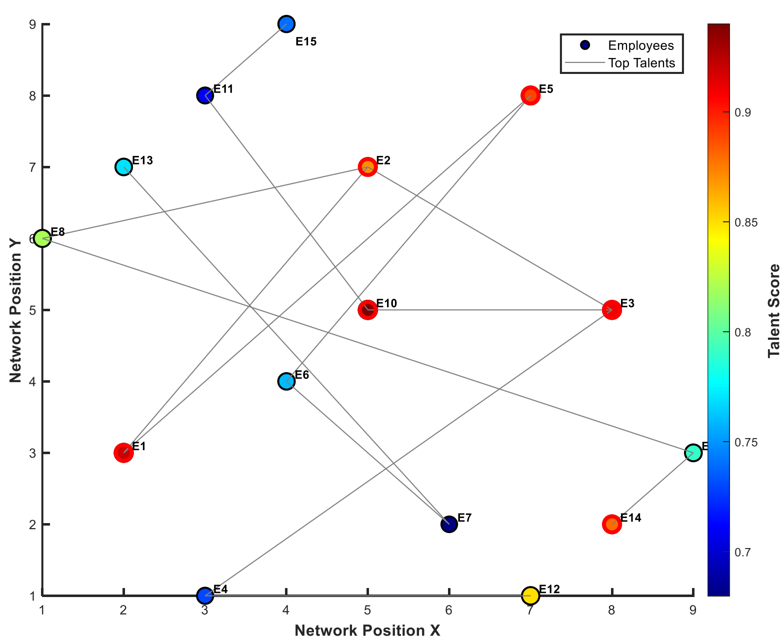

Figure 5. Enterprise talent network analysis and identification results

4.4 Analysis of results

The results of the experiment show that the suggested heterogeneous information network approach has a number of significant advantages over current approaches, especially when it comes to identifying intricate organizational relationships that conventional approaches are unable to. The framework performs exceptionally well in a variety of organizational settings. The sensitivity to data quality issues, especially in skill assessment accuracy, the potential introduction of bias through historical performance data that may not reflect future potential in rapidly evolving technological environments, and the difficulty in scaling computational complexity for organizations with extremely large employee networks exceeding 10,000 individuals are notable limitations (Table 3).

Table 3. Feature importance analysis and meta-path effectiveness

Feature Category	Importance Score	Top Contributing Features	Meta-path Pattern	Effectiveness
Structural Features	0.347	Betweenness Centrality (0.156)	Employee-Employee	High
		PageRank Score (0.142)	Direct Collaboration	High
		Clustering Coefficient (0.049)	Community Structure	Medium
Meta-path Features	0.423	E-P-E Collaboration (0.187)	Employee-Project-Employee	Very High
		E-D-E Cross-functional (0.134)	Employee-Dept-Employee	High
		E-S-E Skill Similarity (0.102)	Employee-Skill-Employee	High
Attribute Features	0.230	Performance History (0.098)	Individual Metrics	Medium
		Skill Diversity Score (0.087)	Competency Breadth	Medium
		Project Impact Rating (0.045)	Contribution Measures	Medium

5. Conclusion and future work

By means of advanced modeling of multi-relational organizational structures, this study presents a novel heterogeneous information network framework that provides interpretable insights into complex talent dynamics within contemporary enterprises and significantly improves enterprise talent identification. The thorough experimental analysis shows that, in comparison to current methods, combining structural network features, meta-path-based relationship patterns, and attribute-driven performance indicators through heterogeneous graph attention mechanisms results in higher accuracy when identifying high-potential employees. The wider impact on the field of HR analytics includes the establishment of heterogeneous information networks as a fundamental paradigm for modeling complex relationships in the workplace. This could revolutionize the way organizations approach succession planning, talent management, and strategic human capital optimization by utilizing data-driven methodologies that recognize the interconnectedness of contemporary collaborative work environments.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically concerning authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere in any language.

Data availability statement

The manuscript contains all the data. However, more data will be available upon request from the corresponding author.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] Tusquellas, N., Palau, R., & Santiago, R. (2024). Analysis of the potential of artificial intelligence for professional development and talent management: A systematic literature review. International Journal of Information Management Data Insights, 4(2), 100288.

[2] Khemani, B., Patil, S., Kotecha, K., & Tanwar, S. (2024). A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. Journal of Big Data, 11(1), 18.

[3] Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., ... & Sun, M. (2020). Graph neural networks: A review of methods and applications. AI open, 1, 57-81.

[4] Qin, C., Zhang, L., Cheng, Y., Zha, R., Shen, D., Zhang, Q., ... & Xiong, H. (2025). A comprehensive survey of artificial intelligence techniques for talent analytics. Proceedings of the IEEE. DOI: 10.1109/JPROC.2025.3572744

[5] França, T. J. F., São Mamede, H., Barroso, J. M. P., & Dos Santos, V. M. P. D. (2023). Artificial intelligence applied to potential assessment and talent identification in an organisational context. Heliyon, 9(4).

[6] Ekuma, K. (2024). Artificial intelligence and automation in human resource development: A systematic review. Human Resource Development Review, 23(2), 199-229.

[7] Outemzabet, L., Gaud, N., Bertaux, A., Nicolle, C., Gerart, S., & Vachenc, S. (2024). Harnessing heterogeneous information networks: A systematic literature review. Computer Science Review, 52, 100633.

[8] Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., & Yu, P. S. (2019, May). Heterogeneous graph attention network. In The world wide web conference (pp. 2022-2032).

[9] Fu, X., Zhang, J., Meng, Z., & King, I. (2020, April). Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. In Proceedings of the web conference 2020 (pp. 2331-2341).

[10] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems, 32(1), 4-24.