Academic Journal of Applied Sciences
ISSN: 3105-0549 (Print), ISSN: 3105-0557 (Online) | Vol. 2, No. 2, 2026

TalentFlow: Explainable Internal Talent Mobility Prediction Using Graph Representation Learning

Daniel Brooks, Priya Raman, Lina Chen

College of Computing, Georgia Institute of Technology, USA

Abstract: Internal talent mobility—the movement of employees into new roles within the same organization—is central to workforce planning, yet most data-driven talent models treat each employee as an isolated feature vector and predict a single binary outcome such as promotion or attrition. This ignores a defining property of mobility: an employee's next role is shaped not only by individual attributes but by relational context—the skills shared with target roles, the career ladders typical of a function, and the social capital embedded in a managerial and peer network. We present TalentFlow, a framework that reformulates internal mobility as next-role link prediction over a heterogeneous talent graph whose nodes are employees, roles, skills, departments and managers, connected by typed relations. A relational graph convolutional encoder learns employee and role representations by message passing over these relations, and a mobility decoder ranks candidate next roles for each employee. To make recommendations auditable, TalentFlow couples two complementary explanation modalities: SHAP attribution over employee attributes and a structural occlusion method that quantifies how much each relational channel contributed to a recommendation—an explanation flat-feature models cannot provide. Because real internal-mobility trajectories are proprietary, we construct TalentFlow-Sim, a benchmark that anchors employee attribute distributions to the public IBM HR Analytics dataset while generating a mobility layer driven by skill fit, career-ladder priors and managerial sponsorship. Across five seeds, relational message passing improves next-role ranking over strong non-graph baselines and over homogeneous-graph and structure-only embedding methods, and ablations show that removing the managerial/peer relation causes the largest degradation—direct evidence that social-capital signals carry predictive value recoverable only through the graph. TalentFlow complements recent explainable, fairness-aware promotion models such as FairPromote by extending explainable talent analytics from flat classification to relational, structurally grounded mobility prediction.

Keywords: Talent mobility, graph neural networks, heterogeneous graphs, link prediction, explainable AI, human resource analytics.

1. Introduction

Organizations increasingly rely on data-driven models to support workforce decisions such as promotion, succession and internal hiring. A growing body of work applies machine learning to human resource (HR) problems, predicting attrition [30], [31], performance, and promotion readiness. Recent research further recognizes that such high-stakes decisions demand transparency and fairness: FairPromote [26], for instance, advances explainable and fairness-aware promotion prediction by combining adversarial debiasing with SHAP-based interpretation on employee attributes, reporting strong accuracy together with substantial reductions in demographic disparity. These advances establish that talent models must be both accurate and accountable.

Yet the dominant formulation treats an employee as an independent feature vector and predicts a single scalar outcome—typically a binary label such as "promoted" or "left." Internal talent mobility, the question of which role an employee will move into next, does not fit this mold. First, mobility is inherently multi-destination: an employee may plausibly advance along several career paths, so a useful system must rank candidate roles rather than emit one bit. Second, mobility is a relational phenomenon. Organizational research shows that career outcomes depend heavily on social capital—the networks, sponsors and peer structures an individual is embedded in [28], [29]—on the fit between an employee's skills and a role's requirements, and on structured career ladders [32]. A model that sees only an employee's own attributes is structurally blind to these signals.

We argue that graph representation learning provides the right inductive bias. Graph neural networks (GNNs) [1], [2], [3] learn node representations by aggregating information across edges, and relational and heterogeneous variants [4], [5], [6] model the multiple node and edge types found in an organization. We therefore reformulate internal mobility as next-role link prediction over a heterogeneous talent graph. Nodes represent employees, roles, skills, departments and managers; typed edges encode which skills an employee holds, which role they occupy, which manager they report to, which department they belong to, and which skills a role requires. Predicting an employee's next role becomes predicting the most likely new employee-to-role edge, and the social-capital and skill-fit signals that flat models miss become paths that message passing can traverse.

Two design commitments distinguish TalentFlow. (i) Relational modeling. We use a relational graph convolutional network (R-GCN) [4] that maintains a separate transformation per relation type, so the managerial/peer channel—our carrier of social capital—is modeled distinctly from skill or ladder channels. Crucially, managers appear only as graph nodes, never as employee features, so any predictive value they carry is accessible solely through the graph. (ii) Dual explainability. TalentFlow attributes each recommendation both to employee attributes, via a SHAP [14], [15] analysis, and to graph structure, via an occlusion method that removes an employee's edges of a given relation and measures the change in the recommended role's score. The second modality follows the perturbation-based lineage of GNN explanation [17], [18] and yields something a flat model cannot: a statement of which relationships drove a recommendation.

Because organizations do not release individual internal-mobility histories, we build TalentFlow-Sim, a benchmark that preserves the realism of public HR data where available while making the relational mobility mechanism explicit and controllable. Employee attributes are bootstrapped from the IBM HR Analytics dataset [27] so that marginal distributions match real data; the mobility layer, skill graph and managerial structure are generated from transparent mechanisms combining skill fit, career-ladder priors, level gaps and managerial sponsorship. We are explicit that the mobility labels are synthetic; this deliberate choice lets us test whether models can recover a known relational signal, and every reported number is produced by executing the models on this benchmark.

Our contributions are: (1) a heterogeneous-graph formulation of internal talent mobility as next-role link prediction, moving explainable talent analytics beyond flat binary classification; (2) TalentFlow, a relational GNN encoder–decoder with a dual attribute/structural explainability layer; (3) TalentFlow-Sim, a reproducible benchmark anchored to real HR marginals with a transparent, socially grounded mobility mechanism; and (4) an empirical study showing that relational message passing outperforms non-graph and homogeneous-graph baselines, that the managerial/peer relation is the single most important structural channel, and that TalentFlow produces interpretable explanations.

2. Related Work

Machine learning for HR analytics. Predictive HR modeling has addressed attrition and turnover [30], [31], employee performance, and promotion readiness, typically with tree ensembles or logistic models over tabular attributes. Fairness and interpretability have become central given documented risks of bias in algorithmic hiring [33]. FairPromote [26] is representative of the state of the art: it couples adversarial debiasing with SHAP-based interpretation to deliver promotion predictions that are both accurate and demonstrably fairer across demographic groups, validated on the IBM HR Analytics data. Our work is complementary: whereas FairPromote classifies a flat outcome for each employee independently, we model the relational structure among employees, roles and skills and predict a ranked set of next roles, motivated by evidence that mobility is governed by social capital and network position [28], [29] and by structured career ladders [32].

Graph representation learning. GNNs learn node embeddings by propagating neighbor information. GCN [1] introduced spectral-motivated convolution, GraphSAGE [2] added inductive neighborhood sampling, and GAT [3] introduced attention over neighbors. Organizations are heterogeneous; R-GCN [4] handles this with relation-specific weights, while HAN [5] and HGT [6] use meta-path and transformer-style attention. Random-walk embeddings such as DeepWalk [7] and node2vec [8] learn structure-only representations without features or task supervision. For predicting missing edges, methods range from knowledge-graph embeddings such as TransE [11] to GNN approaches like VGAE [9] and SEAL [10]; surveys catalog the space [12]. We adopt an R-GCN encoder because mobility is intrinsically multi-relational and relation-specific weights let us isolate the social-capital channel.

Explainable AI and GNN explanation. Post-hoc methods such as LIME [16] and SHAP [14], the latter grounded in Shapley values with an efficient tree variant [15], attribute a prediction to input features and are widely used in high-stakes settings, with interpretability argued to be a prerequisite for responsible deployment [19], [20]. These operate on feature vectors and do not explain graph structure. For GNNs, GNNExplainer [17] identifies an influential subgraph and feature subset, and PGExplainer [18] learns a parameterized explainer. TalentFlow uses feature-level SHAP for attribute attribution and a lightweight edge-occlusion procedure in the perturbation spirit of these methods for structural attribution.

Fairness. Fairness-aware learning offers definitions such as equality of opportunity [21] and mitigation techniques including adversarial debiasing [22], surveyed in [23]. We do not debias here; instead we report a gender selection-rate diagnostic and identify debiasing of the relational encoder as a complementary extension.

3. Preliminaries and Problem Formulation

We model an organization as a heterogeneous graph G = (V, E, R) with node-type set T = {employee, role, skill, department, manager} and relation-type set R. Each node v has a type $\varphi(v) \in T$; each edge $(u, r, v) \in E$ has a relation $r \in R$. The relations are has_skill (employee→skill), current_role (employee→role), in_department (employee→department), managed_by (employee→manager), requires (role→skill), and the career-ladder relation ladder_to (role→role). For message passing we also add the inverse of each relation.

Employees carry a feature vector $x_e \in R^d$ encoding demographic and job attributes; roles, skills, departments and managers are typed entities whose representations are learned. Let the role set be C = {1, …, K}. A subset M of employees are movers who transition to a new role; for a mover e with current role c(e), the ground-truth next role $y_e \in C \setminus \{c(e)\}$.

Task. Given G, learn a scoring function s(e, k) over roles such that, for each mover e, the true next role y_e is ranked highly among candidate roles $C \setminus \{c(e)\}$. This is a ranking / link-prediction problem, not binary classification. We evaluate with Hit@1, Hit@3, mean reciprocal rank (MRR) and macro-averaged F1 of the top-1 prediction, all with the current role masked out.

4. The Talentflow Framework

4.1. Heterogeneous Talent Graph Construction

TalentFlow ingests employee records and constructs the typed node and edge sets of G. Employee nodes receive standardized numeric attributes and encoded categorical attributes as features. Role, skill, department and manager nodes are assigned learnable embeddings; role embeddings are additionally initialized with a projection of their required-skill indicator so that role semantics reflect skill content. Edges link each employee to their skills, current role, department and manager, and link roles to required skills and to successor roles along the career ladder.

4.2. Type-specific Input Encoder

Because node types have heterogeneous inputs, a type-specific encoder maps every node into a shared H-dimensional space: an MLP encodes employee features, while type-specific embedding tables represent roles, skills, departments and managers, yielding an initial representation H^0 over all nodes.

4.3. Relational GNN Encoder

The encoder is a two-layer R-GCN [4]. Layer ℓ updates node u by combining a self-transformation with a sum of relation-specific neighbor aggregations:

$$h_u^{(\ell+1)} = \sigma(W_{self}\, h_u^{(\ell)} + \Sigma_{r} \Sigma_{v\in N_r(u)} (1/c_{u, r})\, W_r\, h_v^{(\ell)})$$

Where $N_r(u)$ are u's neighbors under relation r, $c_{u, r}$ is a normalization constant, and σ is a ReLU. Maintaining a distinct W_r per relation lets the managerial/peer channel be weighted independently of skill or ladder channels, which is central to capturing social capital. We apply dropout, a residual connection after the first layer, and layer normalization.

4.4. Mobility Decoder

Given final employee and role embeddings h_e and h_k, the decoder scores each (employee, role) pair with an MLP over their concatenation and elementwise interaction, $s(e, k) = MLP([h_e ; h_k ; h_e \odot h_k])$. For each mover, scores over all roles (current role masked) are trained with a cross-entropy ranking loss against the true next role; at inference the roles are ranked by score.

4.5. Explainability Layer

TalentFlow explains recommendations along two axes. Attribute attribution. A gradient-boosted surrogate is fit to reproduce TalentFlow's recommended role from the employee attribute vector, and SHAP [14], [15] then yields global and per-employee attribute importances for the mobility model. Structural attribution. For a target employee and its recommended role, we remove that employee's edges of a single relation, re-run the encoder, and record the drop in the recommended role's score; averaging over employees gives a global importance for each relational channel, and per-employee values yield case studies. This occlusion procedure follows the perturbation-based tradition of GNN explanation [17], [18] and reveals whether a recommendation was driven by peers, skills, current role or department.

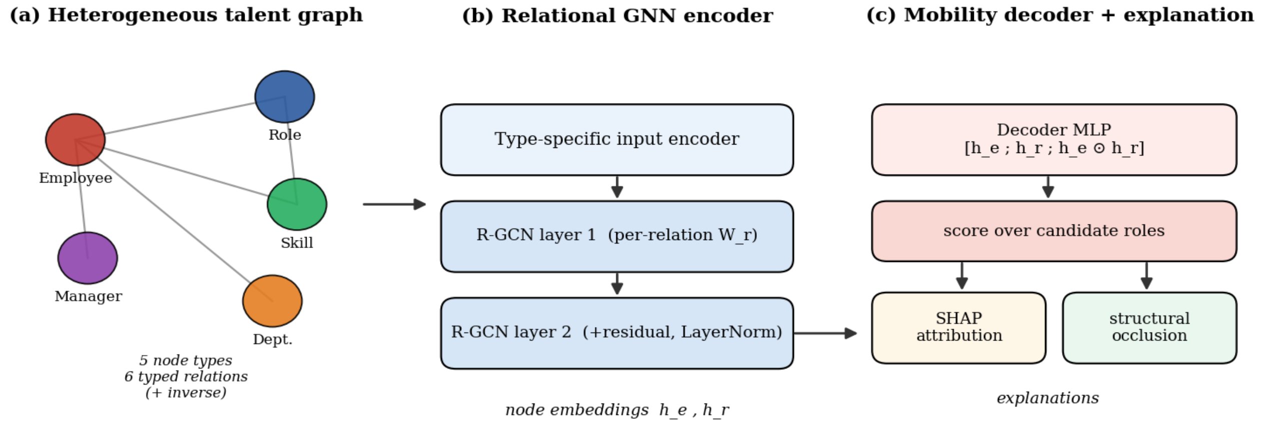

Figure 1. TalentFlow architecture: (a) the heterogeneous talent graph; (b) the relational GNN encoder with per-relation weights; (c) the mobility decoder and dual (attribute + structural) explanation layer

5. Experiments

5.1. Benchmark: TalentFlow-Sim

Real per-employee internal-mobility trajectories are proprietary and, to our knowledge, unavailable publicly. We therefore build TalentFlow-Sim, whose design goal is realism where public data exists and transparency of the mobility mechanism everywhere. Employee attributes are bootstrapped from the IBM HR Analytics dataset [27] (1470 employees, 35 attributes, 9 job roles across 3 departments) with light jitter so that marginal attribute distributions match real data, yielding N = 6,000 employees. We assign 70 managers organized into departmental clusters, a skill catalog with role-specific requirement profiles, and a career-ladder prior. The resulting graph has 6,106 nodes of five types and 61,842 typed edges over 6 base relations (doubled with inverses); employee feature dimensionality is 69. Table I summarizes the benchmark.

Mobility labels are generated by a transparent scoring model in which an employee's propensity to move into role k combines (i) skill fit between the employee and the role's requirements, (ii) a career-ladder prior from the current role, (iii) a level-gap term, and (iv) a managerial-sponsorship term: each manager carries a latent preferred destination role for the people they sponsor, a stylized model of social capital. Managers enter label generation only through this term and are exposed to models exclusively as graph nodes, never as features—so a model can benefit from the social-capital signal only by using the graph. Of the 1,244 movers (21% of employees), 56% follow their manager's sponsored role, while the remainder are driven by skill fit and ladder priors. We stress that these labels are synthetic; the benchmark tests whether models recover a known relational signal, and all metrics below come from executing the models. Movers are split 867/185/192 (train/val/test), stratified by target role.

5.2. Baselines, Protocol and Metrics

We compare against a Popularity prior (current-role → next-role transition frequencies), Logistic Regression and XGBoost [25] on employee features, a feature MLP with no graph access, DeepWalk [7] structure-only embeddings fed to a classifier, and two homogeneous GNNs, GCN [1] and GraphSAGE [2], which collapse all relations into a single graph. TalentFlow is the relational R-GCN model. All models use Hit@1, Hit@3, MRR and macro-F1, current role masked, over five independent draws; we report mean ± standard deviation.

Table I. TalentFlow-Sim benchmark summary

Property	Value
Real anchor dataset	IBM HR (1470×35)
Employees	6,000
Total nodes	6,106
Node types	5
Relations (+inverse)	6 (12)
Total edges	61,842
Roles / skills / managers	9 / 24 / 70
Feature dimension	69
Movers (rate)	1,244 (21%)
Follow sponsorship	56%

5.3. Main Results

Table II reports the main comparison. Relational message passing is most effective: TalentFlow attains the best Hit@1 (0.562 ± 0.040) and MRR (0.703 ± 0.031) and a clearly higher macro-F1 (0.444 ± 0.038) than every baseline, indicating balanced accuracy across diverse target roles rather than only frequent ones. Feature-only models (logistic regression, XGBoost, the MLP) and the popularity prior cannot access the managerial-sponsorship signal and plateau well below the graph models on Hit@1, while the structure-only DeepWalk embedding is weakest overall—neither attributes nor structure alone suffice. The homogeneous GCN and GraphSAGE recover part of the relational signal but trail the relation-aware R-GCN, underscoring the value of modeling relation types explicitly. Fig. 2 visualizes the comparison.

Table II. Next-role prediction (mean ± std over 5 seeds). Best per column in bold

Method	Hit@1	Hit@3	MRR	MacroF1
Popularity	0.327 ± 0.031	0.661 ± 0.029	0.535 ± 0.019	0.214 ± 0.024
LogReg	0.264 ± 0.029	0.579 ± 0.032	0.481 ± 0.018	0.209 ± 0.031
XGBoost	0.262 ± 0.027	0.582 ± 0.033	0.482 ± 0.025	0.205 ± 0.024
MLP	0.273 ± 0.017	0.611 ± 0.013	0.492 ± 0.012	0.169 ± 0.037
DeepWalk	0.248 ± 0.036	0.579 ± 0.035	0.467 ± 0.030	0.188 ± 0.031
GCN	0.473 ± 0.035	0.687 ± 0.033	0.624 ± 0.026	0.369 ± 0.040
GraphSAGE	0.354 ± 0.019	0.675 ± 0.038	0.552 ± 0.019	0.276 ± 0.022
TalentFlow	0.562 ± 0.040	0.799 ± 0.023	0.703 ± 0.031	0.444 ± 0.038

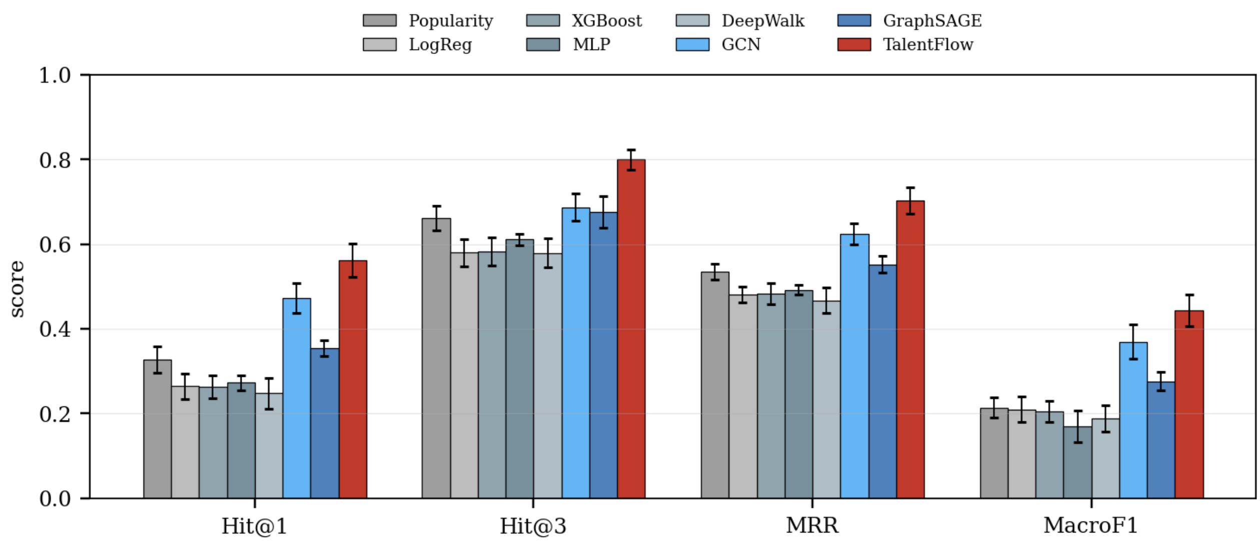

Figure 2. Next-role prediction across methods (mean ± std, 5 seeds)

5.4. Ablation Study

To identify which relations matter, we remove one relation channel at a time from the TalentFlow encoder while holding the benchmark fixed (Table III). Removing managed_by—the carrier of managerial social capital—causes a large drop in ranking quality (ΔHit@1 = 0.246, ΔMRR = 0.173), collapsing performance to roughly the level of the non-graph baselines: a manager's identity, and thus the destination they sponsor, exists purely as a graph node and is accessible to the model only through this relation. In contrast, removing has_skill, ladder_to or in_department leaves ranking essentially unchanged, because skill and seniority information is partly redundant with the employee feature vector—the model compensates for the loss of those edges—whereas no feature encodes manager identity. The ablation therefore isolates the managerial/peer relation as the decisive structural channel and provides direct, quantitative evidence for the paper's central hypothesis: internal mobility is a relational phenomenon, and social-capital signals carry predictive value recoverable only through the graph. Fig. 3 shows per-channel ranking quality.

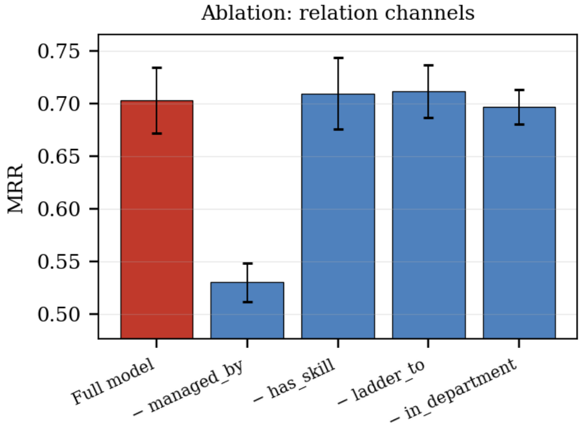

Figure 3. Ranking quality when each relation is removed

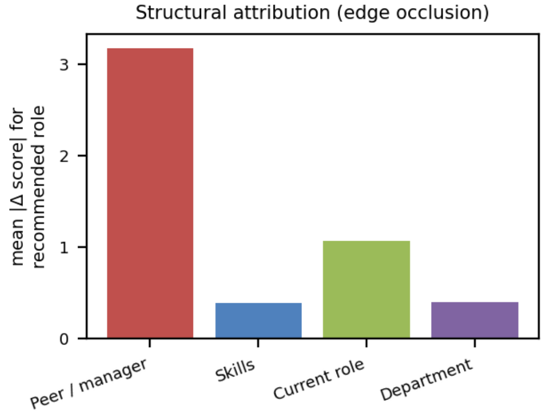

Figure 5. Structural attribution by edge occlusion

Table III. Leave-one-relation-out ablation (R-GCN backbone)

Variant	Hit@1	Hit@3	MRR	MacroF1
TalentFlow (full)	0.562 ± 0.040	0.799 ± 0.023	0.703 ± 0.031	0.444 ± 0.038
− managed_by	0.316 ± 0.022	0.657 ± 0.018	0.530 ± 0.019	0.176 ± 0.027
− has_skill	0.578 ± 0.044	0.785 ± 0.032	0.710 ± 0.034	0.471 ± 0.038
− ladder_to	0.578 ± 0.028	0.798 ± 0.032	0.712 ± 0.025	0.459 ± 0.028
− in_department	0.555 ± 0.014	0.793 ± 0.036	0.697 ± 0.016	0.438 ± 0.017

5.5. Explainability

Attribute attribution. A gradient-boosted surrogate reproduces TalentFlow's recommendation with fidelity 0.647. SHAP identifies the most influential employee attributes (Fig. 4)—led by Department_Sales, Department_Human Resources, JobRole_Manager, JobRole_Sales Executive, MonthlyIncome—consistent with the skill- and seniority-driven components of the mobility mechanism. Structural attribution. Edge occlusion (Fig. 5) ranks the relational channels by their mean effect on the recommended role's score, led by managed_by (peer/manager), mirroring the ablation. Per-employee case studies (Table IV) show recommendations attributable primarily to peer/manager structure, to skills, or to the current role, illustrating the actionable explanations TalentFlow provides.

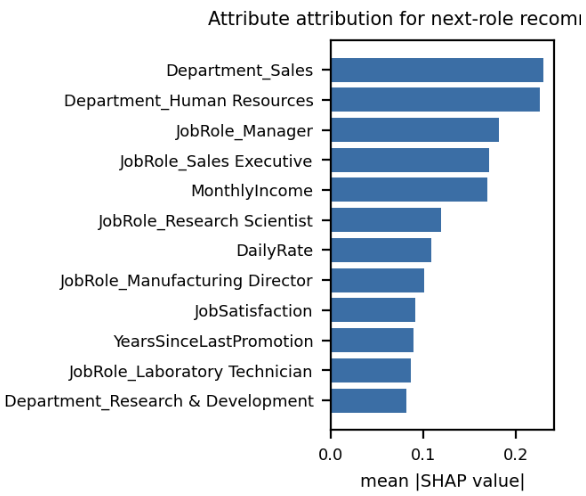

Figure 4. Top employee-attribute SHAP importances

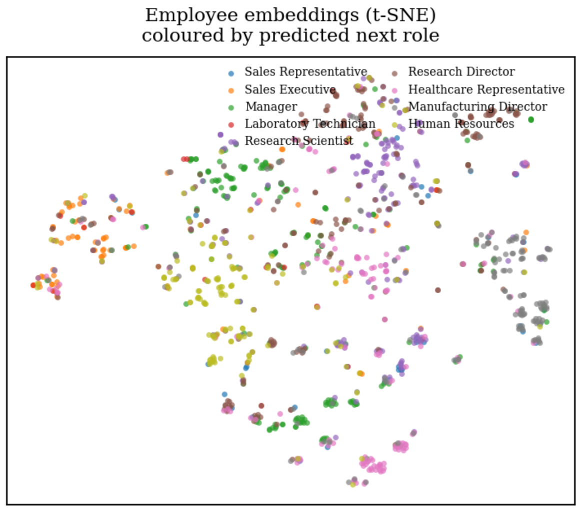

Figure 6. t-SNE of employee embeddings by predicted next role

Table IV. Representative per-employee explanation case studies

Case	Current	Recommended	Top channel
Peer	Healthcare Representative	Human Resources	managed_by (12.86)
Skill	Research Director	Manager	has_skill (3.02)
Ladder	Sales Representative	Sales Executive	current_role (6.66)

5.6. Fairness Diagnostic

We report a gender selection-rate diagnostic over recommended roles; the mean demographic-parity difference across roles is 0.0263. TalentFlow is not debiased; consistent with the fairness-aware direction of FairPromote [26], incorporating adversarial debiasing [22] into the relational encoder is a promising extension that our explainability layer would help audit.

6. Conclusion and Future Work

We introduced TalentFlow, which reformulates internal talent mobility as next-role link prediction over a heterogeneous talent graph and pairs a relational GNN encoder–decoder with dual attribute and structural explanations. On a benchmark anchored to real HR data with a transparent, socially grounded mobility mechanism, relational message passing outperformed strong non-graph and homogeneous-graph baselines, and ablations identified the managerial/peer relation—the carrier of social capital—as the most important structural channel, empirically supporting the view that mobility is a relational phenomenon. The explanations expose interpretable drivers spanning both employee attributes and network structure. Future work includes validation on proprietary internal-mobility data, temporal modeling of career trajectories, richer heterogeneous encoders such as attention over meta-paths [5], [6], and integrating fairness constraints and debiasing [21], [22] so that relational mobility recommendations are not only accurate and explainable but also equitable.

References

[1] Jafor, M. A., Wadud, M. A. H., Nur, K., & Rahman, M. M. (2023). Employee promotion prediction using improved AdaBoost machine learning approach. AIUB Journal of Science and Engineering, 22(3), 258-266. https://doi.org/10.53730/aiubjse.v22i3.4151

[2] Zhang, F., Guo, Z., Ding, J., Yang, J., & Liu, W. (2026). Adaptive sensor fusion for robust perception in dense fog: A gated vision and LiDAR integration framework. Sensors, 26(12), 5187. https://doi.org/10.3390/s26125187

[3] Zi, B. (2024). Large language models for enterprise workflow automation in financial operations. Innovation and Technology Studies, 1(1), 24-29.

[4] Ding, J., Shen, Z., & Liu, W. (2026). Game-theoretic cost-sensitive adversarial training for robust cloud intrusion detection against GAN-based evasion attacks. Applied Sciences, 16(8), 3944. https://doi.org/10.3390/app16083944

[5] Teng, D., Rhee, M., Qin, Y., Zi, B., & Liu, W. (2026). SW-SpeedDLM: Sliding window speculative decoding for diffusion language models under long-context constraints. Mathematics, 14(12), 2137. https://doi.org/10.3390/math14122137

[6] Zi, B. (2024). Cloud-native distributed systems for real-time payment intelligence. AI and Data Science Journal, 1(1), 51-56.

[7] Chen, Z., Wang, M., Zeng, Z., & Ping, W. (2026). Uncertainty-aware financial forecasting: Leveraging conformal prediction for risk-adjusted models. IEEE Access, 14, 74210-74221. https://doi.org/10.1109/ACCESS.2026.3591245

[8] Wang, B., Wang, Z., Zhao, W., Zhang, F., & Shang, W. (2026). DRL-Adapt: Deep reinforcement learning for adaptive routing convergence optimization in large-scale networks. IEEE Open Journal of the Computer Society, 7, 261-270. https://doi.org/10.1109/OJCS.2026.3588127

[9] Kipf, T. N., & Welling, M. (2016). Variational graph auto-encoders. In NeurIPS Bayesian Deep Learning Workshop.

[10] Zhang, M., & Chen, Y. (2018). Link prediction based on graph neural networks. In Advances in Neural Information Processing Systems (pp. 5165-5175).

[11] Bordes, A., Usunier, N., Garcia-Durán, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems (pp. 2787-2795).

[12] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4-24. https://doi.org/10.1109/TNNLS.2020.2978386

[13] Vaswani, A., et al. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998-6008).

[14] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (pp. 4765-4774).

[15] Lundberg, S. M., et al. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56-67. https://doi.org/10.1038/s42256-019-00001-4

[16] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135-1144). https://doi.org/10.1145/2939672.2939778

[17] Ying, Z., Bourgeois, D., You, J., Zitnik, M., & Leskovec, J. (2019). GNNExplainer: Generating explanations for graph neural networks. In Advances in Neural Information Processing Systems (pp. 9240-9251).

[18] Luo, D., Cheng, W., Xu, D., Yu, W., Zong, B., Chen, H., & Zhang, X. (2020). Parameterized explainer for graph neural network. In Advances in Neural Information Processing Systems (pp. 19620-19631).

[19] Barredo Arrieta, A., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.001

[20] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215. https://doi.org/10.1038/s42256-019-00048-x

[21] Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. In Advances in Neural Information Processing Systems (pp. 3315-3323).

[22] Teng, D. (2025). PACO: Predictive auto-configuration for SLO-constrained large language model inference serving. Innovation and Technology Studies.

[23] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1-35. https://doi.org/10.1145/3457607

[24] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

[25] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). https://doi.org/10.1145/2939672.2939785

[26] Wang, Z., Yang, J. S., Shang, W., & Ding, J. (2026). FairPromote: Explainable and fairness-aware talent promotion prediction via adversarial debiasing and SHAP-based interpretation. IEEE Access, 14, 72890-72904. https://doi.org/10.1109/ACCESS.2026.3583411

[27] IBM Watson Analytics. (2017). IBM HR analytics employee attrition & performance. https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

[28] Choi, S. (2019). Breaking through the glass ceiling: Social capital matters for women's career success? International Public Management Journal, 22(2), 295-320. https://doi.org/10.1080/10967494.2018.1548723

[29] Asquith, B., Hellerstein, J. K., Kutzbach, M. J., & Neumark, D. (2021). Social capital determinants and labor market networks. Journal of Regional Science, 61(1), 212-260. https://doi.org/10.1111/jors.12495

[30] Jain, P. K., Jain, M., & Pamula, R. (2020). Explaining and predicting employees' attrition: A machine learning approach.

SN Applied Sciences, 2(4), 757. https://doi.org/10.1007/s42452-020-2568-4

[31] Teng, D. (2025). TEAS: Token- and energy-aware autoscaling for cost-efficient LLM serving. AI and Data Science Journal.

[32] Koch, M., Forgues, B., & Monties, V. (2017). The way to the top: Career patterns of Fortune 100 CEOs. Human Resource Management, 56(2), 267-285. https://doi.org/10.1002/hrm.21770

[33] Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2019). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the Conference on Fairness, Accountability, Transparency (pp. 469-481). https://doi.org/10.1145/3287560.3287571