PACIS 2023 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

7-8-2023

# Acqui-hiring or Acqui-quitting: Post-M&A Turnover Prediction via a Dual-fit Model

Denghui Zhang
*Rutgers Business School*, denghui.zhang@rutgers.edu

Hao Zhong
*ESCP Business School*, hzhong@escp.eu

Jingyuan Yang
*George Mason University*, jyang53@gmu.edu

# Acqui-hiring or Acqui-quitting: Data-driven Post-M&A Turnover Prediction via a Dual-fit Model

*Completed Research Paper*

**Denghui Zhang**
School of Business
Stevens Institute of Technology
Hoboken, New Jersey, United States
dzhang42@stevens.edu

**Hao Zhong**
Information & Operations Management
ESCP Business School
Paris, France
hzhong@escp.eu

**Jingyuan Yang**
Donald G. Costello School of Business
George Mason University
Fairfax, Virginia, United States
jyang53@gmu.edu

## Abstract

*Gaining highly skilled human capital is one of the key motivations for mergers and acquisitions (M&A), particularly in knowledge-intensive sectors such as the technology industry. However, the inherent cultural differences and organizational misalignments during the integration process can lead to significant tensions and a high rate of talent turnover, which may ultimately result in integration failure. Hence, it is crucial for organizations to proactively anticipate and manage the potential effects of such events on employee turnover. The predominant perspective in existing literature focuses on the dyadic relationship between merging firms while a few other studies recognize the fit between employees and the firm. However, there has been a lack of endeavor to unify these two factors into a coherent framework. In this paper, we propose a novel data-driven neural network approach to predict the talent turnover trend during the post-M&A phase, by considering the compatibility between the merging companies as a key factor. Specifically, drawing on organizational theories, we develop a dual-fit heterogeneous graph neural network with 1) Organization to Organization (O-O) fit, which captures the relatedness and similarity at the firm level, and 2) Person to Organization (P-O) fit, which represents the compatibility and cultural closeness at the employee level. By leveraging this framework, we can effectively integrate multi-sourced, heterogeneous data to gain a more nuanced understanding of the compatibility between firm pairs. Our proposed approach is evaluated on a large-scale real-world dataset and benchmarked against state-of-the-art methods. Experimental results demonstrate the superiority of our approach in predicting talent turnover trends during the post-M&A phase. Our approach also offers interpretable results and provides valuable insights for organizations seeking to manage talent effectively during and after M&A events.*

**Keywords:** turnover prediction, mergers and acquisitions, graph neural networks

## Introduction

Over the past few decades, mergers and acquisitions (M&A) have become one of the major strategies for businesses to grow and expand market shares. Driven by an essential need for growth, waves of M&A transactions have reached a historical high, particularly in the technology industry. For example, U.S. mergers and acquisitions sustained a vigorous pace in the first quarter of 2019, with 900 transactions and a total market valuation of $79.5 billion, a 35 percent increase year over year (YoY)[1]. In tech companies, *acqui-hiring* is a trending M&A-based hiring strategy to effectively reinforce the talent pool and boost enterprise value[2] (Chen et al. 2021; Kim 2020). Evidences have shown that tech giants, such as Apple and Google, are leading the way in the AI talent-driven acquisition race in 2016–2020 (GlobalData 2021). To gain and sustain financial benefits of M&A transactions (e.g. revenue and market share gains), companies should also ensure that they have proper management practices and right workforce targets to maintain the acquired intangible assets, i.e., talents. Unfortunately, recent studies (Kim 2020) revealed a notable *acqui-quitting* trend, i.e., 33% of acquired workers quit their jobs within the first year of their employers being acquired. Upon the announcement or speculation of M&A deals, the ensuing integration process can trigger significant transformations and disturbances to the organizational and cultural landscape, leading to profound impacts on employees (Lee and Pennings 1996). In order to maintain productivity and secure sustainable benefits, it is imperative to understand and foresee potential post-M&A talent departures.

Numerous studies have investigated M&A activities from various aspects, including financial performance outcomes, pre-merger firm-level compatibility, organizational culture, and ex-ante M&A experiences, and M&A integration process (Das and Kapil 2012; Trichterborn et al. 2016). While financial performance analysis has received considerable attention in M&A studies (Thanos and Papadakis 2012), post-M&A employee turnover is still under-explored. Only a handful of recent studies started focusing on M&A-related employee turnover by understanding employee attitudes through individual-level primary survey data (Kyei-Poku and Miller 2013). There is a complex mechanism behind employee turnovers, whose driving factors include M&A deal characteristics, managerial effects, organizational compatibility, and cultural fit (Bauer and Matzler 2014; Kim 2020).

To tackle the post-M&A employee turnover prediction problem, we encounter several unique challenges. First, existing literature emphasizes on firm-level characteristics; however, we argue that the compatibility between employees and the acquirer firm cannot be overlooked. Indeed, the M&A turnover involves three primary entities, namely, the acquirer firm, the acquiree firm, and the acquired employees. In order to gain a holistic view of the impact of M&A events on employee turnover, it is essential to take into account not only the acquirer-to-acquiree compatibility, but also employee-to-acquirer fit. Second, studying this tripartite relationship naturally requires a diverse set of data that comprehensively describes all three entities involved. However, such a collection of data is typically large-scale, heterogeneous, and often unstructured, such as company profile descriptions. A conscientious approach to data preparation is needed to effectively leverage the vast amount of unstructured data available and extract meaningful information. Lastly, traditional classification models cannot properly handle the complexity of the three-way relationship along with the heterogeneous and unstructured data. More sophisticated machine learning models are needed to effectively integrate and unleash the full potential of the comprehensive data.

To address the aforementioned challenges, our paper proposes a novel graph neural network-based method to examine the "fit" among these three parties and understand their impacts on employee turnover. In particular, we propose a *Dual-fit* model: an *Organization to Organization fit (O-O fit)* as the measure of firm-level compatibility and complementarity and a *Person to Organization fit (P-O fit)* as the "fit" measure between the acquired employees and the acquirer. Our focus here is whether we could effectively predict the impact of M&A on acquired employee turnover escalation, measured as the difference between pre-M&A vs. post-M&A turnover rates. Common variables used for O-O fit are firm-level characteristics, such as industry/sector, geographic location, and company size and age, to characterize a macro M&A changing environment. While for P-O fit, most literature considered primary data from surveys and interviews capturing employees' subjective perceptions. Only a restricted number of studies have investigated the dual-fit model and performed a fine-grained analysis that integrates a range of diverse data sources.

To this end, we obtain a large-scale heterogeneous dataset consisting of over 2,500 M&A transactions sourced from *Crunchbase*, along with employment profiles of over 806K professionals from *LinkedIn*, which empowers us to conduct an in-depth and fine-grained analysis encompassing a diverse set of factors. Next, we perform data preprocessing and feature engineering to restructure the heterogeneous data and extract relevant patterns by leveraging text data related to company profiles and employee job records. We propose a novel approach to extract complex hidden relationships by introducing a Dual-fit model induced Heterogeneous Graph Neural Network (DHGNN). This model allows for a fine-grained analysis of turnover likelihood among various types

[1]https://www.aerotek.com/en/insights/power-through-m-and-a-disruption-with-a-strong-talent-strategy
[2]https://www.business-sale.com/insights/for-buyers/acquihiring-ma-strategy-to-boost-talent-pool-and-enterprise-value-221601

| | Variables | Remarks | References | Included |
|---|---|---|---|---|
| M&A Transaction | Deal characteristics | Deal attitude, and acquisition premium, etc. | King et al. 2021 | ✗ |
| | Integration process | Integration timeline, depth etc. | King et al. 2021 | ✗ |
| M&A Firm-level | Business category | Market, Industry keywords | King et al. 2021; Shi et al. 2016 | ✓ |
| | Geo-location | City, State, Country | King et al. 2021; Shi et al. 2016 | ✓ |
| | Investors | Investors who have invested the firm | Shi et al. 2016 | ✓ |
| | Top management team | Founders, executives, board members in the firm | Krug et al. 2014 | ✓ |
| Employee Turnover | Job titles | Job function and responsibility terms | Joseph et al. 2007; Steigenberger and Mirc 2020 | ✓ |
| | Job skills | Associated job skill terms | Hang et al. 2022 | ✓ |
| | Education background | Attended schools | Hom et al. 2017; Mobley et al. 1979 | ✓ |
| | Past employment records | Previous employers and job records | Liu et al. 2012 | ✓ |
| | Demographics | Gender, age, and marital status | Mobley et al. 1979 | ✗ |
| | Pay and Promotion | Salary and career development | Joseph et al. 2007; March and Simon 1993 | ✗ |
| | Job performance | Job performance evaluation and assessment | Hom et al. 2017 | ✗ |
| | Organizational Structure | Intra company hierarchical structure | Sun et al. 2019 | ✗ |
| | Social relationship | Work related social interactions, teamwork | Mitchell and Lee 2001; Teng et al. 2019 | ✗ |

**Table 1. Factors Discussed in M&A and Turnover Literature**

of employees. It also provides informative node feature representations of three-way relationships that can reveal rich semantic and structural patterns that are often left undiscovered by traditional classification models or homogeneous graph models. Through extensive experiments and ablation studies on real-world data, we demonstrate that our proposed framework outperforms state-of-the-art benchmarking methods, confirming its superior prediction performance. Overall, our DHGNN model presents a promising approach to understanding complex relationships in employee turnover and can provide valuable insights for effective HR management.

Our research first contributes to the broad M&A literature. To the best of our knowledge, little research has focused on post-M&A employee turnover prediction induced by a dual-fit model. Our research puts a spotlight on this issue and investigates the underlying compatibility and three-way relationship among the acquirer, the acquiree, and the employees. Besides, our research further contributes to the IS and data science literature by demonstrating the power of graph neural network (GNN) models in coping with the post-M&A turnover prediction problem. We tackle this problem by proposing a novel heterogeneous GNN model that captures both firm-level compatibility and employee-firm-level fit. Unlike traditional turnover prediction methods that concern single fit, our dual-fit design can provide a holistic understanding of the complex relationship among the relevant contributing factors by effectively transforming heterogeneous data into rich graph node embedding representations. The proposed techniques can be adopted by researchers and practitioners for other business problems in various business scenarios. Finally, our research has important practical values. Executives and analytics managers can readily adopt our framework to evaluate the potential turnover escalations of various types of employees.

# Literature Review

Our study naturally relates to research on M&A studies, especially from the perspective of employee turnover. Meanwhile, our research is also relevant to strategic management literature as organizational strategic fit and person-organization fit play an important role in employee turnover. On top of these two, we also investigate existing talent turnover prediction models and discuss common variables and machine learning techniques.

## *Mergers and Acquisitions (M&A) and Employee Turnover*

M&As are referred to the events of combining/merging two independent firms into one single entity. There are different types of M&A, including horizontal, vertical, product/market extension, and unrelated. Each M&A type is driven by a distinctive motive, yet all aim for the benefits that arise from the integration of the merging companies (Lukic 2020). Extant literature aimed at understanding the M&A events by studying motivations, the dyadic organizational relatedness or differences, and the M&A effects on performance, especially on financial outcomes (e.g. stock price, profitability, and return on investment) or productivity (e.g. patents) (Lee et al. 2022; Narayanan et al. 2019). Various factors affecting post-acquisition performance have been studied (see Table 1 for details and references), including M&A transaction-related attributes such as deal characteristics and integration process, and the firm-level characteristics in a dyad such as industry, geography, received historical investments as well as human-related factors (e.g. top management team).

Although scholars have primarily focused on the financial ramifications of M&A, there exists a dearth of research attention regarding the effects of M&A on employee-centered outcomes. In practice, it may be partially attributed to the difficulties in collecting the attitude and behavior data from employees throughout the M&A

process. However, it is very crucial to understand the employee side of M&A outcomes to evaluate the deal success, especially for talent-driven acquisitions. Only a few studies starting to analyze the link between employee turnover and M&A performance by exploring various turnover-related factors such as culture, management, and poor motivation. For example, a survey of employees in a Canadian financial institution showed that merger satisfaction and commitment are vital for reducing the post-merger turnover rate (Kyei-Poku and Miller 2013), while Krug et al. 2014 more focused on the top management turnover analysis. Nonetheless, our understanding of post-M&A employee turnover outcomes still remains limited. Our study aims to fill this gap in the literature by incorporating data on occupation, skills, and educational background from publicly available job records. These variables offer valuable insights into the person-organization fit (P-O fit) concept, as discussed further in Section , and enable a more comprehensive comprehension of the prospective turnover patterns.

## *Turnover Theory*

There are many important turnover theories in management literature to explain why employees leave their organizations (Hom et al. 2017), such as the *organizational equilibrium theory* (March and Simon 1993), the *unfolding model* (Lee et al. 1999), and *job embeddedness theory* (Mitchell and Lee 2001). Among the seminal models, the theory of organizational equilibrium stands out, highlighting the pivotal role of two key factors: job mobility and individual proclivity towards attrition (March and Simon 1993). Turnover researchers are actively exploring other significant antecedents to explain employee turnover, for example, social relationship (Teng et al. 2019) and the organizational structure (Sun et al. 2019). Recently, Steigenberger and Mirc (2020) emphasized that organizational and especially under-studied occupational identification have a strong influence on employee turnover decisions. We present in Table 1 several salient factors extracted from the extant employee turnover literature, and specify the relevant variables that have been incorporated in our study. Many turnover studies investigate collective turnover, while there are many more voluntary turnover studies over involuntary ones (Hausknecht and Trevor 2011). Also, turnover research has been conducted over various levels (individual, group/unit, and organization) and different industries and professions (e.g. IT professions (Joseph et al. 2007)). In our paper, we focus on analyzing collective turnover at the organizational level, with particular attention to the effects of organizational-level variables and employee group-specific factors on employee attrition.

## *Organizational Compatibility, Complementarity, and Person-Organization Fit*

There are many dimensions when evaluating the match between the target firm and the acquirer firm. One notable stream of strategic management research studies the impact of the fit between two firms on the M&A success (Bauer and Matzler 2014; Homburg and Bucerius 2006).

**Organizational Compatibility and Complementarity.** In the school of strategic management literature, the core concept is that high compatibility (relatedness or similarity) in the management styles and organizational culture can effectively increase value creation and boost synergy realization (Bauer and Matzler 2014). In other words, the similarity has a positive impact on performance and also reduces the potential cultural conflicts during the integration process. A few studies focus on the measure of the similarity of the firm-level characteristics between the acquirer and acquiree, such as the match or compatibility among the company pair as measured in several proximity metrics (Gomes et al. 2013; Shi et al. 2016). On the other hand, strategic complementarity has been widely discussed with the aspects of market complementarity (Kim and Finkelstein 2009), technological complementarity (Makri et al. 2010), or product and resource complementarity (Wang and Zajac 2007). Most of the empirical evidence from the literature found that strategic complementarity has positive influences on cultural fit and the ease of integration (Bauer and Matzler 2014). We adopt the aforementioned mechanisms and focus on measuring the O-O fit gauging the level of similarity/proximity between the firm-level attributes of the acquiring and target companies.

**Person-Organization Fit.** In addition to the firm-level compatibility, we also need to consider the Person-Organization (P-O) fit when considering employee turnover. P-O fit has been defined as “the compatibility between the employee and organization” and described as “a multidimensional construct consisting of three determinants of fit: values, personality, and environment” (O’Reilly III et al. 1991; Westerman and Cyr 2004). It is a strong indicator of the employees’ attraction to the organization as well as the intention to stay within the organization in the future (Kristof-Brown and Guay 2011). However, the changes brought by the M&A may result in employees from the acquired company experiencing a sense of misfit with the acquirer company due to the introduction of new vision, management styles, and potential culture shock (Buono and Bowditch 2003).

| Variable | Count | Variable | Mean | Std | Min | Max |
|---|---|---|---|---|---|---|
| # M&A deals | 2,566 | # M&A deals per acquirer | 1.38 | 1.31 | 1 | 24 |
| # acquirers | 1,861 | # M&A deals per acquiree | 1 | 0 | 1 | 1 |
| # acquirees | 2,566 | # industry keywords per acquirer | 3.73 | 1.89 | 1 | 13 |
| # locations | 947 | # industry keywords per acquiree | 3.3 | 1.63 | 1 | 11 |
| # industry keywords | 48 | # investors per acquirer | 3.53 | 3.19 | 1 | 23 |
| # investors | 3,758 | # investors per acquiree | 3.52 | 2.61 | 1 | 18 |
| # employees | 806,536 | # employees per acquiree | 327.31 | 1,717.94 | 1 | 43,175 |
| # top mgmt members | 17,191 | # top mgmt members per acquiree | 3.73 | 3.8 | 1 | 52 |
| # employee groups | 64 | # employees per employee group | 15,985.91 | 22,447.95 | 44 | 118,910 |
| # job records | 1,212,319 | # employee groups per acquiree | 11.32 | 11.52 | 1 | 60 |
| # job skill terms | 262,791 | # job records per employee | 1.5 | 1.11 | 1 | 29 |
| # schools | 27,113 | # job skill terms per employee | 13.78 | 17.96 | 0 | 104 |

**Table 2. Descriptive Statistics of Our Data Sample**

The majority of previous studies on P-O fit rely on primary data collected through methods such as interviews and surveys that assess job satisfaction and turnover intention (Greenwood et al. 1994). Our work instead will primarily utilize objective measures to analyze the impact of M&A on P-O fit. Specifically, we emphasize the match between the acquired employees and the acquirer firm based on the job position, skills and social factors (alumni) and also embed the fit in a heterogeneous graph to form a dual-fit graph for the M&A company pairs.

In summary, this study aims to incorporate both O-O fit and P-O fit into a heterogeneous graph neural network to provide a more holistic understanding of turnover in the context of M&A within the relationships between acquirers, targets, and employees. Using the constructed comprehensive graph, a variety of factors have been integrated to evaluate the dual fit, including the acquirer company's prior experiences with acquisitions, the degree of industry and business relatedness between the companies, and the composition of the top management team, among others.

### *Graph Neural Networks*

Our methodology relates to the broad literature of Graph Neural Networks (GNN) models. GNN models have recently caught wide and unprecedented attention in data mining and machine learning communities as there are a growing number of applications where data are represented in the forms of graphs/networks. According to the types of nodes and/or edges in a network, GNN models can be classified into *homogeneous network-based models* (only a single type of node and edge) and *heterogeneous network-based models* (with multiple types of nodes and/or edges). As for homogeneous GNNs, convolutional GNNs appeared to be the mainstream and popular models: including GCN (Kipf and Welling 2016) and GAT (Veličković et al. 2018). To cope with more complex heterogeneous networks, heterogeneous GNNs were later developed, which consist of proximity-preserving-based models, message-passing-based models, and relation-learning-based models (Yang et al. 2020). Our model fits into the class of message-passing methods, where representative models include *RGCN* (Schlichtkrull et al. 2018) and *HGT* (Hu et al. 2020).

Our study stands out from previous research in two key ways. First, although commonly used subjective factors are effective in predicting employee turnover decisions, the primary survey data takes a long time to collect and thus is not easily available. Therefore, it is of more practical value to build the prediction model based on the readily available objective variables for both firms and employees. Second, our M&A turnover prediction is under a unique setting of the M&A merging impact and thus a more complex dual-fit among the three parties needs to be considered. Therefore, heterogeneous GNN is preferred to deal with multiple entity types.

## Data and Preliminary Analysis

### *M&A Data Collection*

Our data are collected from two sources. The first is *Crunchbase*, a premier database of startup activities, investments and funding information, founding members and key personnels, mergers, and acquisitions (M&A), and industry trends[3]. *Crunchbase* is well-recognized and has rising potential for economic and managerial research (Butler et al. 2020; Dalle et al. 2017). We rely on this database to gather data on firm demographics, M&A deals, investments, and firms' key members. The second data source is *LinkedIn*, one of the major

[3]https://en.wikipedia.org/wiki/Crunchbase

| Group | Functionality | Responsibility |
|---|---|---|
| 1 | project, technical, engineering, development, support technical, design, technology, production | manager, manager senior, supervisor, manager regional, director management, associate manager |
| 2 | software, systems, solutions, enterprise, qa, security, solution, tech, development software, digital | lead, representative, director senior, leader, senior, associate senior, designer senior, lead senior |
| 3 | business, business development, strategic, corporate, business operations, strategy | vp, agent, rep, operator, controller, ii, co, client, owner, clerk |
| 4 | customer service, service, customer, customer support, care customer, customer success | editor, writer, producer, artist, senior writer, editor senior, associate producer, editor writer |
| 5 | data, test, research, human resources, field, clinical, medical, health, information technology, healthcare | engineer, engineer senior, developer, technician, engineer lead, engineer principal, engineer staff |
| 6 | operations, product, support, team, program, quality, system, hr, network, channel | specialist, analyst, consultant, analyst senior, consultant senior, advisor, expert, consultant principal |
| 7 | supply chain, retail, commercial, electrical, store, products, storage, travel, purchasing | director, president vice, executive, head, executive senior, associate director, director regional |
| 8 | finance, financial, market, contract, compliance, accounting, accounts, credit, accounts strategic | intern, associate, assistant, trainer, recruiter, instructor, trainee, internship, student, generalist |

**Table 3. Common Job Terms in *FUN* and *RES* Groups.**

professional networking platforms. As of March 2022, it has over 810 million registered users from over 200 countries and territories worldwide[4]. It has been used as the major data source in various research projects to help understand career paths (Lappas 2020) and labor markets (Liu et al. 2020b). Likewise, we acquired thorough profiles of workers (e.g., skills, career history, and educational background) from *LinkedIn*.

Our data collection process starts with sampling M&A deals. And we restrict our focus to the M&A deals completed after the year 2000 (inclusive) and the acquirees founded after 1990 (inclusive) with headquarters in the United States. To avoid any discrepancies, the two data sources are then linked by ensuring the exact matching of firm names. Individuals' profiles are largely retained for the sake of capturing more prominent career mobility patterns. Table 2 showcases our final data sample's descriptive statistics. There are in total 2,566 M&A deals and 1,861 acquirers. The majority of M&A transactions occurred between 2010 and 2018. The firms in our data sample are distributed in 947 different US cities, most of which are located in "New York", "San Francisco", and "Austin". Within our dataset, the firms are associated with 48 unique industry keywords that define their market sectors. In contrast to standard industry codes such as SIC code[5], our industry keywords offer a more detailed and specific categorization of firms. We have also accumulated substantial employment-related data of acquiree firms, including 800K employees and over 1.2M job records. Each acquiree has an average of 327 employees and each employee has average of 1.5 job records. In addition, we have collected around 17K profiles belonging to top management members (including executives, founders, and board members) from both organizations engaged in M&A activities. Our dataset comprises a total of 260,000 unique skill terms, with each employee being associated with an average of 13.78 ones. Please refer to Table 2 for additional details.

## *Employee Group (EMG)*

We note again that the main objective of our research is to understand the impact of M&A transactions on the turnover rate of various types of employees. We therefore carefully categorize employees according to their occupational genres, which include *functionality (FUN)* and *responsibility (RES)*. Given a job title, we extract *FUN* and *RES* using the method proposed by Liu et al. (2020a). The procedure for constructing employee groups is outlined as follows. We begin by classifying each job title according to its relationship to the two genres. We acquire job term embeddings by applying a pre-trained word embedding model, GloVe (Pennington et al. 2014), which encodes semantic meaning into vectors. Next, mini-batch K-Means clustering is employed on the two genres of job term embeddings to generate genre-specific clusters for the job titles. To determine the proper number of clusters for two genres respectively, we employ the Elbow method (Thorndike 1953) and choose $K^{FUN} = 8$ and $K^{RES} = 8$. Table 3 lists most common job terms in each *FUN* and *RES* groups. Our final employee groups are determined by the joint genre-specific cluster IDs. For example, suppose that firm $i$ belongs to $C_i^{FUN} = 3$ and $C_i^{RES} = 5$, its employee group is defined as $C_i = \{3, 5\}$. With such clustering outcomes, we observe an average of 11.32 employee groups per acquiree and 16K employees per employee group, as indicated in Table 2.

[4]https://about.linkedin.com/
[5]https://siccode.com/sic-code-lookup-directory

| Variable | Freq | Turnover Rate (Post.-Pre.) | Turnover Escalation % |
|---|---|---|---|
| Overall | 21,237 | 0.012 | 12.5% |
| Location (different country) | 4,726 | -0.007 | 10.2% |
| Location (same country, different state) | 12,507 | 0.016 | 12.7% |
| Location (same state, different city) | 2,601 | 0.027 | 14.8% |
| Location (same city) | 1,403 | 0.011 | 14.3% |
| Industry keywords (no overlap) | 9,623 | 0.007 | 12.7% |
| Industry keywords (one in common) | 9,101 | 0.019 | 12.7% |
| Industry keywords (two in common) | 2,382 | 0.005 | 11.5% |
| Industry keywords (three or more in common) | 125 | 0.026 | 9.6 % |
| Top management (no overlap) | 20,208 | 0.012 | 12.6% |
| Top management (at least one in common) | 1,029 | 0.008 | 10.4% |
| Employees' skills (no overlap) | 8,151 | 0.009 | 11.5% |
| Employees' skills (at least one in common) | 13,086 | 0.014 | 13.1% |
| Employees' schools (no overlap) | 11,960 | 0.009 | 11.4% |
| Employees' schools (at least one in common) | 9,277 | 0.016 | 14.0% |
| Firms' investors (no overlap) | 20,778 | 0.011 | 12.4% |
| Firms' investors (at least one in common) | 459 | 0.041 | 15.9% |

Note: the reduced number of observations in certain groups is attributed to the missing data.

**Table 4. Summary of Employees' Turnover Escalation**

### *Target Variable*

Here, we elaborate on how the target variable *Turnover Escalation* is defined in our study. We define *Turnover Escalation* as the significant rise of the turnover rate in the post-M&A period compared with that in the pre-M&A period. Regarding any M&A event, we formally define three distinct periods as shown in Figure 1. To ensure sufficient data records, we set an 18-month observation window for pre- and post-M&A periods. Meanwhile, we embed a 3-month buffer period prior to the M&A event date to eliminate possible turnover data contamination due to internal message leakage. Note that *Turnover Escalation* is computed at the level of *Acquirer - Acquiree - Employee_Group* (**ACR-ACE-EMG**). Given pre- and post-M&A periods, we first aggregate the number of turnovers in EMG group $k$ for any M&A event between acquirer $i$ and acquiree $j$, i.e., $N_{ijk}^{pre}$ and $N_{ijk}^{post}$. With the total number of employees $N_{jk}^{total}$ in EMG group $k$ of acquiree $j$, we calculate the difference of turnover rates between pre-M&A and post-M&A periods as:

$$\Delta R_{ijk} = R_{ijk}^{post} - R_{ijk}^{pre} = \frac{N_{ijk}^{post}}{N_{jk}^{total}} - \frac{N_{ijk}^{pre}}{N_{jk}^{total}}. \tag{1}$$

Lastly, we perform proportion tests of $\Delta R_{ijk} > 0$ and label *Turnover Escalation* $= 1$ if the difference is significant at the level of significance 0.01, otherwise *Turnover Escalation* $= 0$.

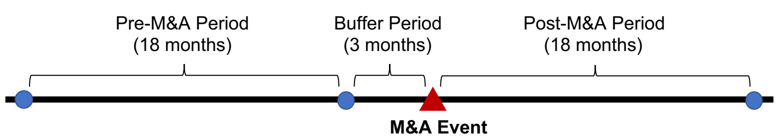


**Figure 1. Definition of Three Periods for M&A Events**

### *Model-free Evidences*

Prior to our model development, we discuss some model-free evidences from our data. Basically, we investigate how pre- and pose-M&A turnover rates ($R^{pre}$ and $R^{post}$) vary while considering specific types of M&A events. Table 4 presents the M&A events grouped by diverse attributes of firms, employees, and investors. For each group, we show the **ACR-ACE-EMG** triplet count, the turnover rate difference (Post.- Pre.), and the rate of turnover escalation. Our analysis reveals that the overall post-M&A turnover rate surpasses the pre-M&A turnover rate, which is consistent with the conclusions drawn in a prior study (Kim 2020). As a frequently discussed factor in the existing literature (Shi et al. 2016), we explore the impact of geographic proximity between acquirers and acquirees by examining whether both entities are located in the same country, state, or city. We also consider business proximity as another variable of interest (Tuch and O'Sullivan 2007), which involves measuring the number of shared industry keywords between the two firms. Subsequently, we analyze three employee-related features shared between acquirers and acquirees, specifically, the presence of shared

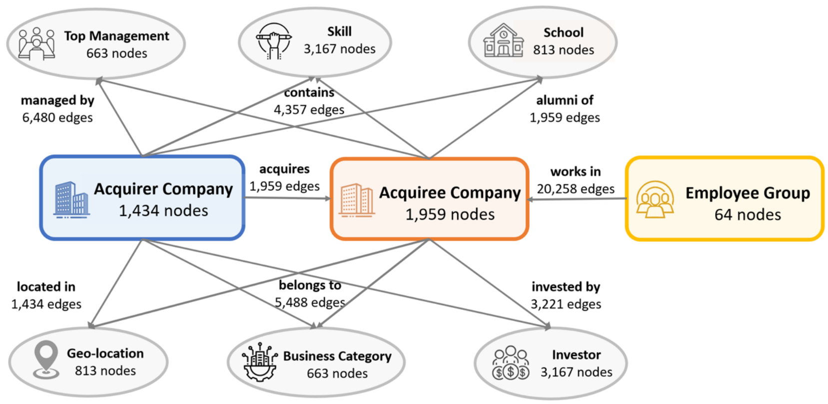


**Figure 2. The Meta Graph of Heterogeneous Organization-Employee Graph**

top management personnel, common job skills, and attended educational institutions by their employees. Finally, we explore the presence of common investors between the acquirers and acquirees in the past. We will conduct a further investigation of these proximity measures in our experimental analysis.

# Methodology

## *Problem Definition*

Our objective is to predict post-M&A turnover trend by framing the task as a **binary classification problem**. We focus on a particular group (type) of employees $\mathrm{EMG}_k$ within an acquiree company $\mathrm{ACE}_j$, assuming that it will be acquired or merged by an acquirer company $\mathrm{ACR}_i$. Specifically, we aim to determine whether the turnover rate $\mathrm{R}_{\mathrm{EMG}_k}$ of this employee group in the acquired company will experience a significant increase following the M&A announcement. To achieve this, we take as input an Acquirer-Acquiree-EmployeeGroup triplet ($\mathbf{ACR}_i$, $\mathbf{ACE}_j$, $\mathbf{EMG}_k$) of features and output a binary variable $\boldsymbol{y} \in \{0, 1\}$, where a value of $1$ indicates a significant increase in turnover rate and $0$ indicates no significant change. Section *Target Variable* explains how we obtain the target binary training labels for $\boldsymbol{y}$.

## *Heterogeneous Organization-Employee Graph (HOEG)*

With the flexibility and expressive power of graphs as well as the heterogeneity of post-M&A data, we create a Heterogeneous Organization-Employee Graph (HOEG) by converting our data into heterogeneous graph data.

**Definition 1** ***Heterogeneous Graph.*** *The heterogeneous graph is a type of graph consisting of different node types and link types. Let $G < V, E >$ denote a graph, where $V$ denotes the node set, $E$ denotes the edge set. Then $G < V, E >$ is heterogeneous when it contains a list of nodes types $V = \{V_1, V_2, ..., V_N\}$, where $N > 1$. Each type $V_i$ contains $n_i$ nodes: $\{t_{i,1}, t_{i,2}, ..., t_{i,n_i}\}$. Equally, it should also contain a list types of edges $E = \{E_1, E_2, ..., E_M\}$, where each type $E_i$ contains $m_i$ edges: $\{e_{i,1}, e_{i,2}, ..., e_{i,m_i}\}$.*

For the post-M&A turnover trend prediction task, we define a special heterogeneous graph, namely, Heterogeneous Organization-Employee Graph (HOEG), to represent all heterogeneous objects in our data. Figure 2 shows the meta graph of HOEG in which there are three types of **core node**, i.e., {$V_1$=Acquirer Company, $V_2$=Acquiree Company, $V_3$=Employee Group}, six types of **supplementary nodes**, i.e., {$V_4$=Business Category, $V_5$=Geo-location, $V_6$=Investor, $V_7$=Top Management, $V_8$=School, $V_9$=Skill}. The core node types correspond to the input triplet ($\mathrm{ACR}_i$, $\mathrm{ACE}_j$, $\mathrm{EMG}_k$) while the supplementary nodes correspond to objects in attribute columns of the input triplet. In other words, all the valid values in these six attributes are transformed into different types of nodes in the graph, e.g., "Oakland, California, United States" for Geo-location, and "HTML" for Skill. Noted that Figure 2 contains fewer nodes than the original data in Table 2 since we only built the graph on the training set.

## *Dual-fit Heterogeneous Graph Neural Network*

**Overview.** Figure 3 shows an overview of our Dual-fit Heterogeneous Graph Neural Network (DHGNN) for post-M&A turnover trend prediction. Our model mainly consists of two parts: Organization-Organization Fit Network (OOFN, O-O Fit Network) and Person-Organization Fit Network (POFN, P-O Fit Network). Given an input triplet ($\mathrm{ACR}_i$, $\mathrm{ACE}_j$, $\mathrm{EMG}_k$), we first locate the corresponding core nodes as well as surrounding sup-

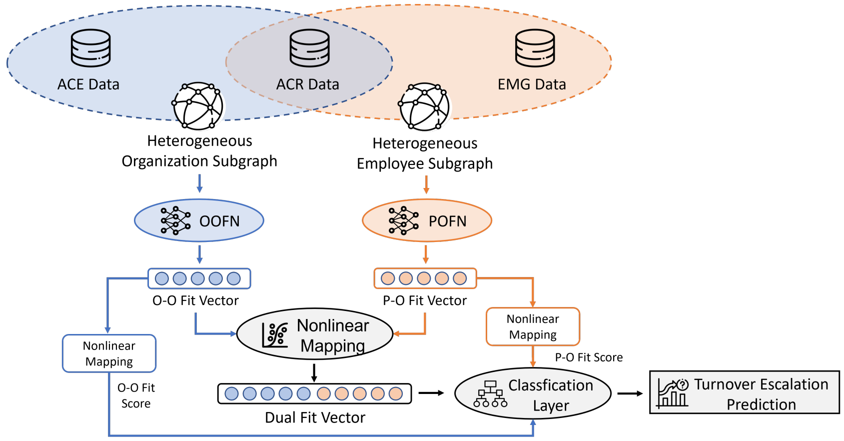


**Figure 3. Dual-fit Heterogeneous GNN for Triplet-based M&A Turnover Prediction**

plementary nodes in HOEG, these nodes automatically constitute two heterogeneous subgraphs, i.e., heterogeneous organization subgraph and employee subgraph. The former subgraph is centered by acquirer, acquiree node $\text{ACR}_i$, $\text{ACE}_j$ while the latter is centered by acquirer node $\text{ACR}_i$, employee group $\text{EMG}_k$. O-O Fit Network (left part in Figure 3) will apply graph convolution operations on $\text{ACR}_i$, $\text{ACE}_j$ nodes to encode the attributes and heterogeneous neighborhood into the hidden representations. Later we further concatenate two hidden vectors of $\text{ACR}_i$, $\text{ACE}_j$ and apply non-linear layer to explicitly generate an O-O Fit score, which aims to model the compatibility and complementarity between the acquirer and acquiree. Similarly, P-O Fit Network (right part) will encode essential information of $\text{ACR}_i$ and $\text{EMG}_k$ into their node representations and generate a P-O Fit score, which models the compatibility between a certain employee group (in acquiree) and the acquirer company. In the last step, we combine P-O Fit and O-O Fit vectors into one unit by non-linear transformations and feed into the classification layer to make the final turnover trend prediction.

## Heterogeneous Message Passing

The fundamental idea of most graph neural networks (GNNs) is Message Passing – aggregating feature information from a node's direct (first-order) neighbors, such as GCN (Kipf and Welling 2016) or GAT (Veličković et al. 2018). The general message passing scheme given a node $x_i$ in the graph is defined as follows:

$$x_i' = \gamma\big(x_i, \rho_{j\in\mathcal{N}}\phi(x_i, x_j)\big) \tag{2}$$

where $\phi$ is the message function, depends on node feature $x_i$ , $x_j$. $\rho_{j\in\mathcal{N}}$ denotes the aggregation function (one can choose sum or average, etc,), $\gamma$ is the update function, i.e., the final transformation to obtain new attributes after aggregating the message. First, each node in the graph computes a message for each of its neighbors. Then each node aggregates the messages it receives using a permutation-invariant function (i.e., the order of the message does not matter). Upon receiving the messages, each node updates its attributes based on its current attributes and the aggregated messages.

Obviously, Message Passing assumes that the graph only contains one type of node and each node only contain one type of feature ($\mathcal{N}$ in Eq.(2) means homogeneous neighbor and $x$ is a homogeneous feature). The assumption is too strong for our setting. Actually, in HOEG, the Acquirer and Acquiree nodes contain numerical content, such as "Company Size" and "Company Age", whereas other nodes only contain categorical content. As a result, we require different feature transformations to handle different types and dimensions of features.

To tackle the issues, we propose a novel Heterogeneous Message Passing method for HOEG. Specifically, we adopt a more flexible assumption that each node type may contain multiple features. Given a node type, let $\mathcal{X}_i = \{x_{i,1}, x_{i,2}, ..., x_{i,n}\}$ denotes the heterogeneous feature set for node $v_i$. We define a new message function that takes heterogeneous features:

$$\phi(v_i, v_j) = \frac{\sum_{x_{j,n}\in\mathcal{X}_j}\left[\overrightarrow{LSTM}\{\mathcal{F}(x_{j,n})\}\bigoplus\overleftarrow{LSTM}\{\mathcal{F}(x_{j,n})\}\right]}{|\mathcal{X}_j|} \tag{3}$$

where $v_i$ is the center node, $v_j$ is one of the neighbor node of $v_i$, $\mathcal{X}_j$ is feature set of $v_j$. As can be seen, we use bi-directional LSTM (bi-LSTM) (Hochreiter and Schmidhuber 1997) to capture "deep" interactions among heterogeneous features and encode them into the same feature space. $\mathcal{F}(\cdot)$ denotes a feature transformer that

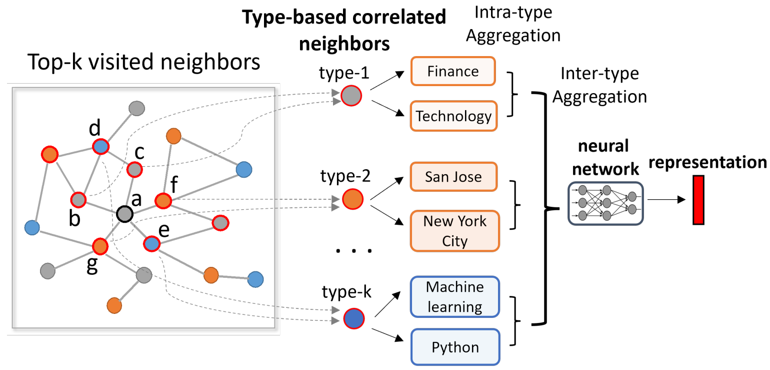


**Figure 4. Heterogeneous Neighbor Nodes Aggregation (Intra-type and Inter-type)**

ensures equal-length input features. $\bigoplus$ denotes vector concatenation.

Next, we define a new aggregation function that aggregates heterogeneous neighbor nodes in two steps: *Intra-type Aggregation* and *Inter-type Aggregation*. Intra-type aggregation first aggregates all neighbor nodes of the same type and then inter-type aggregation combines all different types. For a specific node $v$, we iterate all its neighbors and get a list of $t_n$ types of nodes. For each type $t$, we first perform *intra-type aggregation*:

$$\rho_1^t(v) = \mathcal{G}^t_{v' \in \mathcal{N}_t(v)}\{\phi(v, v')\} \tag{4}$$

where $\mathcal{G}^t\{\cdot\}$ denotes the aggregator for node type $t$, which can be a fully connected network or recurrent neural network. We use a fully connected network as the aggregation function. $\mathcal{N}_t(v)$ denotes the sampled type $t$ neighbors for node $v$, here we adopt a normalized sampling ratio to ensure balanced samples among different types, that is, we use a fixed sampling ratio $r_0$ times $r_t$ the ratio of type $t$ nodes in all graph nodes. $\phi(\cdot)$ denotes the content aggregation function defined previously. For a better understanding, we show an example in Figure 4, where node type 1 represents "Business Category", and two different nodes under this type ("Finance" and "Technology") will be aggregated first following the above process.

Then we perform *inter-type aggregation* to further aggregate the above results of different types,

$$\rho_2(v) = \alpha^{v,v}\phi(v, v) + \sum_t \alpha^{v,t}\rho_1^t(v) \tag{5}$$

where $\phi(v, v)$ denotes the aggregated content embeddings of node $v$ itself. $\alpha(v, t)$ is the learnable attention weights indicating the importance of the corresponding neighbor type $t$ to node $v$, defined as follows:

$$\alpha^{v,t} = \frac{exp\{ReLU(W^\top[\phi(v, v) \bigoplus \rho_1^t])\}}{\sum_{t \in \mathcal{T}(v) \bigcup \{v\}} exp\{ReLU(W^\top[\phi(v, v) \bigoplus \rho_1^t])\}} \tag{6}$$

where $ReLU(\cdot)$ denotes the non-linear function of Rectified Linear Unit, $W^\top$ denotes the attention parameter, $\bigoplus$ denotes concatenation. Lastly, we use the $\rho_2(v)$ to update the embedding of node $v$, in other words, we use the identity function as our updating function $\gamma(\cdot)$. Similarly, in the example of Figure 4, three different node types (type-1, 2, and k represent Business Category, Geo-location, and Skill respectively) will be further aggregated following the above procedure.

Basically, Eq.(3)-(6) constitutes the entire heterogeneous message passing method, which can aggregate heterogeneous neighbor nodes as well as their heterogeneous contents.

## Organization-Organization Fit and Person-Organization Fit

Recall that the input to our model is a triplet of ($\text{ACR}_i$, $\text{ACE}_j$, $\text{EMG}_k$), we apply the O-O fit network and P-O fit network to obtain two fit vectors, which model the M&A deal fitness from the organization perspective and employee perspective, respectively.

**O-O Fit Network and P-O Fit Network.** In O-O fit network, we first apply two layers of the Heterogeneous Message Passing function (Eq.(3)-(6)) on the acquirer node $v_i^{acr}$ and acquiree node $v_j^{ace}$ and their neighbor nodes to generate the aggregated feature embeddings $x_i^{acr}$, $x_j^{ace}$,

$$x_i^{acr} = \rho_2(\rho_2(v_i^{acr})), x_j^{ace} = \rho_2(\rho_2(v_j^{ace})), \tag{7}$$

The two embeddings above contain heterogeneous information such as "location", "business category", "investors", etc. Then we concatenate them and further apply a non-linear transformation layer to obtain the O-O fit vector, i.e., $x_{i,j}^{acr,ace}$, which encodes the combinational information of acquire and acquiree. By later end-to-end training, the vector should learn what combinations of the heterogeneous information in the two firms will result in a good fit that keeps the post-M&A turnover rate stable. For P-O fit, we apply the similar procedures as O-O fit on the acquirer node $v_i^{acr}$ and the employee group node $v_k^{emg}$ to obtain the P-O fit vector.

### *Loss Function and Model Training*

Following the model design in Section *Organization-Organization Fit and Person-Organization Fit*, as the final step, for the input triplet $(\text{ACR}_i, \text{ACE}_j, \text{EMG}_k)$, we use the O-O fit vector $x_{i,j}^{acr,ace}$ and P-O fit vector $x_{i,k}^{acr,emg}$ to generate O-O fit score $s_{oo}$ and P-O fit score $s_{po}$ respectively via fully-connected layers and non-linear activation function (sigmoid). Each score is in the range of $[0, 1]$. Lastly an overall fit score $s_{i,j,k}$ for the input triplet is obtained by averaging O-O fit and P-O fit scores, i.e., $s_{i,j,k} = (s_{oo} + s_{po})/2$. We train the entire model using the cross-entropy as the loss function and adopt Adam optimizer (Kingma and Ba 2015) for mini-batch stochastic gradient descent,

$$\mathcal{L} = -\sum_{i,j,k \in \mathcal{B}} [y_{i,j,k} \cdot \log(1 - s_{i,j,k}) + (1 - y_{i,j,k} \cdot \log(s_{i,j,k}))] \tag{8}$$

where $\mathcal{B}$ denotes one random batch of the entire triplets data, $(i, j, k) \in \mathcal{B}$ stands for the indices for the triplet from the current batch. $y_{i,j,k}$ is the groundtruth label of turnover escalation. To be noted, the fit score $s_{i,j,k}$ has the opposite optimization direction to the turnover escalation prediction variable. In other words, the larger the fit score is (i.e., $s_{i,j,k} \to 1$), the less likely a turnover escalation happen (i.e., $y_{i,j,k} = 0$).

## Experimental Results

We created a comprehensive experimental dataset using the data described in Section *Data and Preliminary Analysis*. Each sample in this dataset follows a specific format: $\{$triplet id | raw attributes | handcrafted features | heterogeneous neighbor nodes | target label$\}$. This format has the advantage of being compatible with a broad range of baseline models, as well as our graph-based model. Different models might exclude specific columns, e.g., conventional ML models only use handcrafted features. We split the entire dataset into training/validation/test sets, in a ratio of 6:2:2, and trained all models using the training set. We fine-tuned the hyperparameters of the models on the validation set (except for the naive models), and computed evaluation metrics on the test set.

### *Baselines, Evaluation Metrics and Implementation Details*

We perform experimental analysis and comparison on different models. First, we consider two mean-based models as our entry baselines. The first mean-based model is **Industry+Age**, in which we compute each acquiree's post-M&A turnover escalation using the corresponding mean estimate of all acquirees in the same industry and with the same firm age. The second mean-based model is denoted as **EMG**, which relies solely on employ group (EMG), i.e., using mean estimates of all M&A deals with the same EMG as the predictions. Second, we also built four conventional ML-based models on top of the hand-crafted features as another set of baseline models, namely, Logistical Regression (**LR**), Support Vector Machine (**SVM**), Decision Tree (**DT**), and Random Forest (**RF**). We intend to demonstrate the best-attainable performance using ML models without embedding-based features. Third, we include three GNN-based models: 1) We downgrade our heterogeneous network into a homogeneous one by ignoring type variation of nodes and edges to train a **GCN** model (Kipf and Welling 2016); 2) As an example of HGNN models, **RGCN** (Schlichtkrull et al. 2018) is incorporated given its popularity and fit to our problem; 3) A more recent heterogeneous graph transformer model **HGT** (Hu et al. 2020) is included which reveals promising performance in large heterogeneous graphs. Lastly, we include our Random Forest model built with embedding-based features (**Embedding+RF**), two downgraded versions of our dual-fit model (**O-O Fit** and **P-O Fit**), and our proposed model **Dual-Fit (DHGNN)**.

To evaluate the prediction performance of all models, we adopted a variety of classification metrics for a comprehensive evaluation: Precision, Recall, F1-score, AUC (Area under the ROC Curve). For Precision, Recall, and F1-score, we use both the Macro and Micro averaging method to calculate them. The difference between Macro and Micro averaging is that Macro averaging gives equal weight to each category while Micro averaging

| | | Precision (%) | | Recall (%) | | F1-score (%) | | AUC (%) |
|---|---|---|---|---|---|---|---|---|
| | | Macro | Micro | Macro | Micro | Macro | Micro | |
| Mean-based models | Industry+Age | 52.75 | 79.78 | 55.99 | 55.96 | 46.93 | 63.07 | 59.84 (+24.4%) |
| | EMG | 51.93 | 79.11 | 54.23 | 52.97 | 44.88 | 60.46 | 56.62 (+31.5%) |
| Conventional ML models | LR | 53.06 | 80.06 | 56.64 | 56.61 | 47.45 | 63.64 | 58.79 (+26.7%) |
| | SVM | 52.57 | 79.64 | 55.61 | 55.65 | 46.64 | 62.82 | 58.05 (+28.3%) |
| | DT | 57.95 | 83.76 | 66.44 | 66.41 | 55.81 | 71.78 | 70.88 (+5.1%) |
| | RF | 58.37 | 84.03 | **67.20** | 67.18 | 56.49 | 72.40 | 72.88 (+2.2%) |
| Existing GNN models | GCN | 68.56 | 83.11 | 58.97 | 85.32 | 59.43 | 82.70 | 72.11 (+3.3%) |
| | RGCN | 68.98 | 83.24 | 59.12 | 85.88 | 59.81 | 83.30 | 72.98 (+2.1%) |
| | HGT | 69.07 | 83.42 | 59.73 | 86.09 | 60.02 | 83.21 | 73.10 (+1.9%) |
| **Our models** | Embedding+RF | 69.50 | 84.58 | 57.80 | 87.10 | 59.04 | 82.76 | 72.85 (+2.3%) |
| | O-O Fit | 69.85 | 83.44 | 59.36 | 86.89 | 60.58 | 83.98 | 73.39 (+1.5%) |
| | P-O Fit | 66.18 | 82.13 | 56.71 | 85.16 | 58.92 | 82.33 | 71.54 (+4.1%) |
| | **Dual-fit (DHGNN)** | **70.45** | **84.94** | 60.32 | **87.67** | **62.40** | **84.94** | **74.49** |

**Table 5. Overall Performances on Post-M&A Turnover Prediction**

gives equal weight to each sample. For example, the micro average precision is the sum of all true positives divided by the sum of all true positives and false positives. The macro average precision is the arithmetic mean of all the precision values for the different classes.

We omit the Accuracy metric since it will be biased by our imbalanced data (90% class 0). To obtain the best performance of our model, we empirically tuned the hyperparameters on the validation set and performed grid search over the following parameter values: learning rate = $\{0.00001, 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.1\}$, batch size = $\{16, 32, 64, 128\}$, initial node feature dimension = $\{64, 128, 256, 512\}$, embedding dimension = $\{64, 128, 256\}$. To ensure robust results, we ran the fine-tuning 10 times and take the average. The optimal hyperparameters for our model are: learning rate $= 0.001$, batch size $= 512$, initial node feature dimension $= 128$, embedding dimension $= 128$. We also performed grid search for baseline models and reported their best performances.

## *Results Analysis*

**Overall prediction performance and ablation study.** Table 5 shows the performance of all models using the default classification threshold (0.5) where bold numbers indicate the best results. Regarding the overall prediction performance on both classes (i.e., turnover escalation and non-escalation), we examined the AUC and F1-score and observed that: our complete model (DHGNN) achieved the best results among all models. This confirms the superiority of our DHGNN model over the mean-based models, conventional ML models as well as existing GNNs. Mean-based models are simple and fast, but do not have strong predictive power, whereas, ML models learn strong patterns that can generalize to test data. The results of our exploratory model (Embedding+RF) improved significantly over conventional ML models, indicating that pre-trained embeddings can serve as feature augmentation, which aligns with our intuition of learning better M&A object embeddings using heterogeneous graphs. Existing GNN models especially RGCN and HGT, which are designed for heterogeneous graphs, achieved the second-best among all baselines, validating the predictive power of heterogeneous graphs. We also conducted the model ablation study to examine each component in our model. We compared our complete dual-fit model with each of our sub-models (O-O fit and P-O fit) and observed that: O-O fit network alone yielded acceptable performance (better than RF+Embedding), whereas, P-O fit alone gave relatively poor results (worse than RF). This indicates O-O fit played a more important role than P-O fit in M&A fit modeling. As only combining them together resulted in better performances, it is evident that both O-O and P-O fit contributes to the superior performance of the entire model.

In addition, to further validate the effectiveness of our proposed heterogeneous graph, we conducted a node-type ablation study to investigate the effects of different nodes. We run the same experiment on our model six more times by removing one type of node in the graph at each round (Table 6). We observed that nodes of {*Category, Skill, Investor*} play relatively more important roles than nodes of {*Top Management, School, Location*}, since the performance drops more after removing the first three types of node. While at the meantime, all of the six node types contribute a bit to our final results as performance will drop once removing any of them.

**Discussions on fit scores.** We continue our discussions on the two fit scores by visualizing their distribution using a *Heat Map* in Figure 5. Each of the two fit scores are binned into 10 equal-width bins (thus 2D squares in the plot). In each square, the color shade indicates the proportion of *Turnover Escalation* cases, i.e., darker color indicates more turnover escalations. A 2D Gaussian filter is applied to increase the smoothness of the distribution. We have several interesting observations from Figure 5. First, there is a clear and smooth

| | | Precision (%) | | Recall (%) | | F1-score (%) | | AUC (%) |
|---|---|---|---|---|---|---|---|---|
| | | Macro | Micro | Macro | Micro | Macro | Micro | |
| DHGNN Ablations | w/o Category | 68.75 | 82.42 | 57.21 | 85.45 | 59.05 | 82.37 | 71.91 |
| | w/o Skill | 68.98 | 82.55 | 57.76 | 85.86 | 59.31 | 82.43 | 72.12 |
| | w/o Investor | 69.18 | 82.83 | 58.72 | 85.89 | 59.97 | 82.94 | 72.38 |
| | w/o Top Management | 69.47 | 82.13 | 59.25 | 86.11 | 59.82 | 82.93 | 73.18 |
| | w/o School | 69.43 | 82.75 | 59.72 | 86.92 | 60.98 | 83.30 | 73.61 |
| | w/o Location | 70.11 | 84.17 | 59.78 | 86.88 | 61.12 | 83.97 | 73.39 |
| | **Full Model** | **70.45** | **84.94** | **60.32** | **87.67** | **62.40** | **84.94** | **74.49** |

**Table 6. Ablation Study of DHGNN Model**

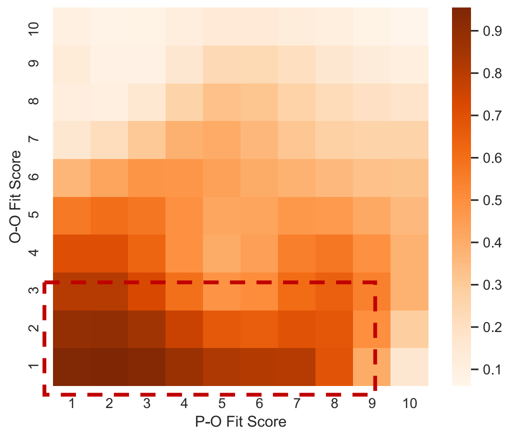


**Figure 5. Overall Distribution of Fit Scores**

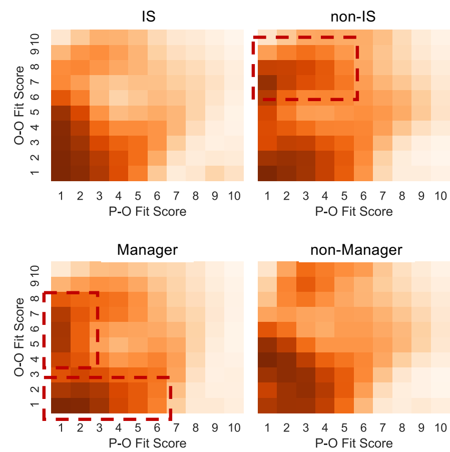


**Figure 6. Fit Scores by EMG**

| **Proximity** | **Elements from Two Firms** | **Calculation** |
|---|---|---|
| Geographic Proximity | Country, State, City | From 1 to 3 based on whether Country, State, or City is the same |
| Business Proximity | Keywords in business category | Sentence embedding cosine similarity |
| Top Management Proximity | Top management team members | Intersection and normalized by sigmoid |
| Skill Proximity | Common skills that employees hold | Intersection and normalized by sigmoid |
| School Proximity | Common schools employees go to | Intersection and normalized by sigmoid |
| Investor Proximity | Historical investors | Intersection and normalized by sigmoid |

**Table 7. Six Proximity Measures and Calculating Methods**

transition of color shades from bottom-left to top-right, which well reflects that our two fit scores largely coincide with true *Turnover Escalations*. Second, most of the darker squares reside in the near-diagonal regions, implying that we should not overlook either of the two fit scores when investigating post-M&A turnovers. Meanwhile, for the darker shades near the bottom (marked by a red box), we can observe that when O-O fit score is low and P-O fit varies from low to high, we may constantly get high turnover escalations, re-affirming our earlier argument that O-O fit potentially contributes more in identifying *Turnover Escalation* cases.

Owing to our model's unique capability of differentiating EMGs, we can investigate the distribution of the two fit scores in greater detail. First, we are interested in the difference between IS (Information Systems) and non-IS related employee groups. Typically, IS jobs require technical expertise, such as programming, data analysis, and system administration, whereas non-IS jobs may require different types of skills, such as communication, problem-solving, and customer service. We thus segment all EMGs into IS-related (FUN #1, #2, #5 groups in Table 3) and non-IS-related groups. Then, we plot the two fit scores for each group separately, as shown at the top of Figure 6. We observe a darker region in the top-left corner of the non-IS plot, which indicates that non-IS employees have a higher chance of quitting if they cannot fit in the new company even in the case of high O-O fit scores. Likewise, we re-segment all EMGs into Manager-related (RES #1, #2, #3, and #7 groups in Table 3) and non-Manager-related groups and show the fit score distribution at the bottom of Figure 6. We find that, in the plot of Manager-related groups, darker squares digress from the diagonal. It may imply that senior-level employees are more likely to quit if single fit scores are too low (i.e., the fit scores are notably unbalanced).

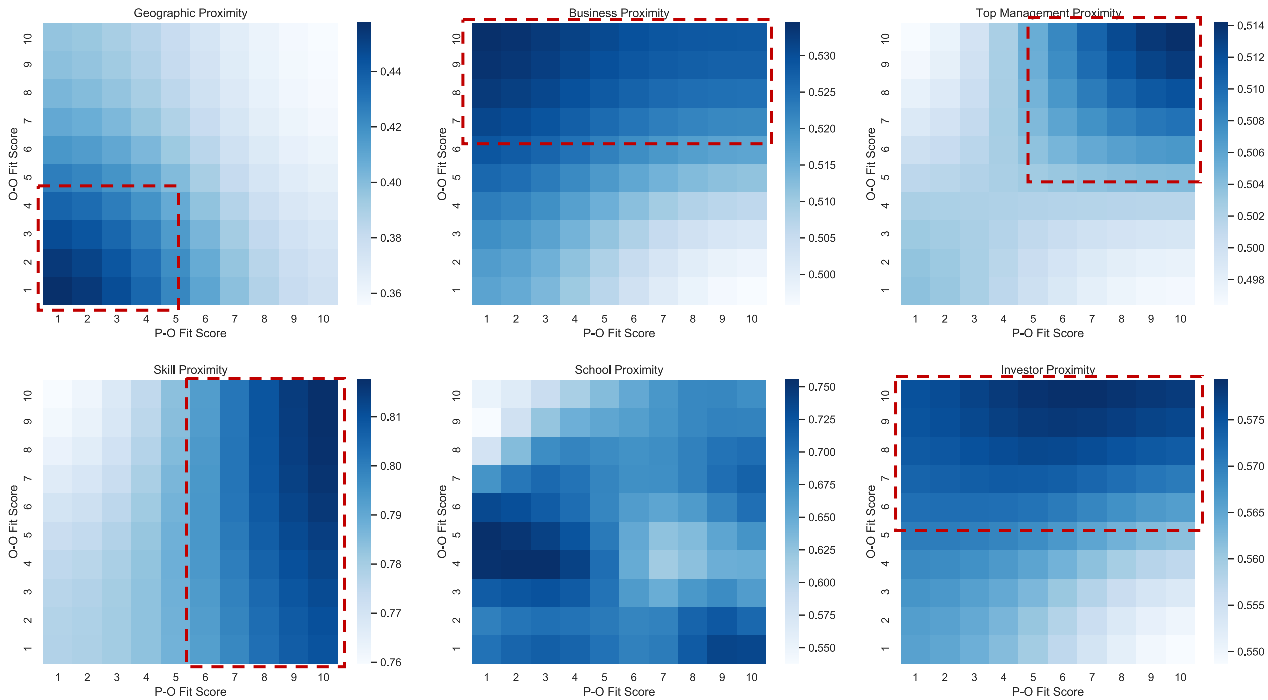


**Figure 7. Distribution of Various Proximity Measures with Fit Scores**

**Various proximity measures and fit scores.** We delve deeper into exploring the relationship between a diverse set of proximity measures and the two fit scores. Specifically, we measure the *Geographic Proximity*, *Business Proximity*, *Top Management Proximity*, *Skill Proximity*, *School Proximity*, and *Investor Proximity* between the acquirer company and acquiree company, and compare these measures with our two fit scores in heatmaps to see their correlations. The calculation of these proximity measures is outlined in Table 7.

As seen in Figure 7, we start by investigating the proximity measures between two companies involved in an M&A transaction. Darker color shades indicate higher proximity scores. In view of geographic proximity, our findings reveal that M&A events tend to occur in two scenarios: 1) when the two companies are far apart (low geographic proximity), it tends to have high fit scores; 2) when the two companies are geographically close, high fit scores are not necessarily a prerequisite. It is rational to assume that the likelihood of successful integration between two organizations may be enhanced by their geographic proximity, as it can facilitate greater interaction and collaboration, thereby potentially improving the outcome of the M&A, irrespective of their level of fit. However, in situations where the two companies are geographically distant, a higher level of dual-fit may be necessary to ensure a successful integration. Additionally, we observe a strong correlation between O-O fit scores and business proximity (i.e., the similarity between the two companies' industry sectors). This observation is reasonable and reinforces the effectiveness of our O-O fit score, as it should partially reveal the degree of similarity between the business operations of the two companies.

We then turn to the proximity measures associated with employees. First, our investigation into top management proximity clearly indicates a positive correlation with high O-O and P-O fit scores. It is arguable that the proximity of top management members, who hold key positions in both organizations, could cultivate a stronger shared company culture, consequently resulting in a higher level of fit. Second, we find that a greater degree of skill proximity is linked to an elevated level of P-O fit across both entities. This finding is well-supported, as employees are predominantly characterized by their job-specific skills. Third, we note that neither the O-O nor the P-O fit scores exhibit a clear correlation with the similarity of attended schools of the employees. This finding is consistent with the outcomes of our earlier ablation study, which indicated that this attribute has a limited contribution to the overall performance of our model. Lastly, higher investor proximity is associated with higher O-O fit scores in comparison to P-O fit scores. It appears to be a logical outcome since firm-level characteristics, as opposed to individual employee-level attributes, tends to be more homogeneous when two companies are under the same investor umbrella.

**Sensitivity analysis on model hyperparameters.** We have also done a sensitivity analysis on multiple hyperparameters that affect our model performance. The performance is relatively stable (with all the AUCs above 70%) despite varying the learning rate, batch size, and embedding dimension, indicating that our model can be easily fine-tuned to achieve optimal performance.

## Conclusion

In this paper, we study how to predict post-M&A turnover escalation at a fine-grained level by considering both the merging firm-level fit and the person-level fit. To the best of our knowledge, this is the first work that tackles this problem by using a heterogeneous graph neural network approach to extract the dual-fit of the three-way relationship among the acquirer, the acquiree, and the employees. Existing work on M&A analysis primarily accentuates the financial ramifications and strategic management procedure, with insufficient attention paid to employees' reaction to the changes brought by M&A events. To effectively capture the intricate relationship among the acquirer-acquiree-employees triplets, we initially procured extensive multi-sourced datasets, which encompass detailed information about the profiles of the M&A firms and the job histories of the acquired employees. We then propose a novel heterogeneous graph neural network approach with a dual-fit design (DHGNN), to extract informative node feature representations of the three-way relationships. This method reveals rich semantic and structural patterns that would remain undetected through conventional classification models or homogeneous graph models. We conducted a comprehensive set of experiments to assess the efficacy of our DHGNN model on the real-world dataset. Our findings indicate that the proposed DHGNN approach outperforms all the benchmarking methods based on classification metrics, thus demonstrating its superior performance. And both the experimental results and ablation study underscore the significance of our dual-fit design and the heterogeneous graph node embeddings in driving the proposed approach's effectiveness.

This study offers a new perspective on acquired workforces, forming a three-way relationship among the acquirer firm, the acquiree firm, and the acquired employees. Our approach bridges the gap in M&A literature by providing a comprehensive understanding and evaluation of post-M&A turnover prediction. We propose an innovative way of predicting post-M&A talent risk based on readily accessible firm data and employee professional profiles, enabling M&A firms to efficiently and effectively predict turnover rates. This can lead to a reduction in turnover rates, improved employee retention, and increased productivity, enhancing overall organizational performance. Furthermore, predicting employee turnover can help organizations make informed decisions about which categories of employees to retain and develop, contributing to the long-term success of the acquired firm. Although this study focuses on the M&A context, the dual-fit GNN framework is transferable to other situations where organizational compatibility and employee coordination are critical. This research could be extended to any business context that involves substantial changes in ownership, control, and strategic direction, necessitating careful planning and execution to achieve success.

## References

Bauer, F. and Matzler, K. 2014. "Antecedents of M&A success: The role of strategic complementarity, cultural fit, and degree and speed of integration". *Strategic Management Journal* (35:2), pp. 269–291.

Buono, A. F. and Bowditch, J. L. 2003. *The human side of mergers and acquisitions: Managing collisions between people, cultures, and organizations*. Beard Books.

Butler, J. S., Garg, R., and Stephens, B. 2020. "Social networks, funding, and regional advantages in technology entrepreneurship: An empirical analysis". *Information Systems Research* (31:1), pp. 198–216.

Chen, D., Gao, H., and Ma, Y. 2021. "Human capital-driven acquisition: evidence from the inevitable disclosure doctrine". *Management Science* (67:8), pp. 4643–4664.

Dalle, J.-M., Den Besten, M., and Menon, C. 2017. "Using Crunchbase for economic and managerial research". *OECD Science, Technology and Industry Working Papers* (2017/08).

Das, A. and Kapil, S. 2012. "Explaining M&A performance: a review of empirical research". en. *Journal of Strategy and Management* (5:3), pp. 284–330.

GlobalData 2021. *Apple the top acquirer of AI companies while other US tech giants also among the forerunners, says GlobalData*. https://www.globaldata.com/apple-top-acquirer-ai-companies-us-tech-giants-also-among-forerunners-says-globaldata/.

Gomes, E., Angwin, D. N., Weber, Y., and Yedidia Tarba, S. 2013. "Critical success factors through the mergers and acquisitions process: revealing pre-and post-M&A connections for improved performance". *Thunderbird International Business Review* (55:1), pp. 13–35.

Greenwood, R., Hinings, C. R., and Brown, J. 1994. "Merging professional service firms". *Organization Science* (5:2), pp. 239–257.

Hang, J., Dong, Z., Zhao, H., Song, X., Wang, P., and Zhu, H. 2022. "Outside in: Market-aware heterogeneous graph neural network for employee turnover prediction". In: *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 353–362.

Hausknecht, J. P. and Trevor, C. O. 2011. "Collective turnover at the group, unit, and organizational levels: Evidence, issues, and implications". *Journal of Management* (37:1), pp. 352–388.

Hochreiter, S. and Schmidhuber, J. 1997. "Long short-term memory". *Neural Comp.* (9:8), pp. 1735–1780.

Hom, P. W., Lee, T. W., Shaw, J. D., and Hausknecht, J. P. 2017. "One hundred years of employee turnover theory and research." *Journal of Applied Psychology* (102:3), p. 530.

Homburg, C. and Bucerius, M. 2006. "Is speed of integration really a success factor of M&A? An analysis of the role of internal and external relatedness". *Strategic Management Journal* (27:4), pp. 347–367.

Hu, Z., Dong, Y., Wang, K., and Sun, Y. 2020. "Heterogeneous graph transformer". In: *Proceedings of The Web Conference 2020*, pp. 2704–2710.

Joseph, D., Ng, K.-Y., Koh, C., and Ang, S. 2007. "Turnover of information technology professionals: A narrative review, meta-analytic structural equation modeling, and model development". *MIS Quarterly* (), pp. 547–577.

Kim, J. D. 2020. *Startup acquisitions as a hiring strategy: worker choice and turnover*. en. SSRN Scholarly Paper ID 3252784. Rochester, NY: Social Science Research Network.

Kim, J.-Y. and Finkelstein, S. 2009. "The effects of strategic and market complementarity on acquisition performance: Evidence from the US commercial banking industry, 1989–2001". *Strategic Management Journal* (30:6), pp. 617–646.

King, D. R., Wang, G., Samimi, M., and Cortes, A. F. 2021. "A meta-analytic integration of acquisition performance prediction". *Journal of Management Studies* (58:5), pp. 1198–1236.

Kingma, D. P. and Ba, J. 2015. "Adam: A Method for Stochastic Optimization". In: *Proceedings of the International Conference on Learning Representations*.

Kipf, T. N. and Welling, M. 2016. "Semi-Supervised Classification with Graph Convolutional Networks". *CoRR* (abs/1609.02907). arXiv: 1609.02907.

Kristof-Brown, A. and Guay, R. P. 2011. "Person–environment fit." In: *APA handbook of industrial and organizational psychology, Vol 3: Maintaining, expanding, and contracting the organization*. American Psychological Association, pp. 3–50.

Krug, J. A., Wright, P., and Kroll, M. J. 2014. "Top management turnover following M&A: Solid research to date but still much to be learned". *Academy of Management Perspectives* (28:2), pp. 147–163.

Kyei-Poku, I. A. and Miller, D. 2013. "Impact of employee merger satisfaction on organizational commitment and turnover intentions: A study of a Canadian financial institution". *International Journal of Management* (30:4), p. 205.

Lappas, T. 2020. "Mining career paths from large resume databases: evidence from IT professionals". *ACM Transactions on Knowledge Discovery from Data (TKDD)* (14:3), pp. 1–38.

Lee, K., Han, K., Animesh, A., and Pinsonneault, A. 2022. "Does IT Matter to Acquisitions? The Impacts of IT Distance on Post-Acquisition Performance". *MIS Quarterly* (46:4), pp. 2261–2288.

Lee, K. and Pennings, J. M. 1996. *Mergers and acquisitions: strategic–organizational fit and outcomes*. Tech. rep. The Wharton School, University of Pennsylvania.

Lee, T. W., Mitchell, T. R., Holtom, B. C., McDaneil, L. S., and Hill, J. W. 1999. "The unfolding model of voluntary turnover: A replication and extension". *Academy of Management Journal* (42:4), pp. 450–462.

Liu, D., Mitchell, T. R., Lee, T. W., Holtom, B. C., and Hinkin, T. R. 2012. "When employees are out of step with coworkers: How job satisfaction trajectory and dispersion influence individual-and unit-level voluntary turnover". *Academy of management journal* (55:6), pp. 1360–1380.

Liu, J., Ng, Y. C., Wood, K. L., and Lim, K. H. 2020a. "IPOD: a large-scale industrial and professional occupation dataset". In: *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing*, pp. 323–328.

Liu, Y., Pant, G., and Sheng, O. R. 2020b. "Predicting labor market competition: leveraging interfirm network and employee skills". *Information Systems Research* (31:4), pp. 1443–1466.

Lukic, A. 2020. "Corporate Mergers and Acquisitions: Longitudinal Consequences for Employee Fit, Satisfaction, and Turnover". In: Arts.

Makri, M., Hitt, M. A., and Lane, P. J. 2010. "Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions". *Strategic Management Journal* (31:6), pp. 602–628.

March, J. G. and Simon, H. A. 1993. *Organizations*. John wiley & sons.

Mitchell, T. R. and Lee, T. W. 2001. "The unfolding model of voluntary turnover and job embeddedness: Foundations for a comprehensive theory of attachment". *Research in Organizational Behavior* (23), pp. 189–246.

Mobley, W. H., Griffeth, R. W., Hand, H. H., and Meglino, B. M. 1979. "Review and conceptual analysis of the employee turnover process." *Psychological bulletin* (86:3), p. 493.

Narayanan, S. et al. 2019. "Analysis of merger & acquisition frameworks from a deal rationale perspective in technology sector". PhD thesis. Massachusetts Institute of Technology.

O'Reilly III, C. A., Chatman, J., and Caldwell, D. F. 1991. "People and organizational culture: A profile comparison approach to assessing person-organization fit". *Academy of Management Journal* (34:3), pp. 487–516.

Pennington, J., Socher, R., and Manning, C. D. 2014. "Glove: Global vectors for word representation". In: *Proceedings of Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543.

Schlichtkrull, M., Kipf, T. N., Bloem, P., Van Den Berg, R., Titov, I., and Welling, M. 2018. "Modeling relational data with graph convolutional networks". In: *European Semantic Web Conference*. Springer, pp. 593–607.

Shi, Z., Lee, G. M., and Whinston, A. B. 2016. "Toward a Better Measure of Business Proximity: Topic Modeling for Industry Intelligence". *MIS Quarterly* (40:4), pp. 1035–1056.

Steigenberger, N. and Mirc, N. 2020. "Should I stay or should I go? Multi-focus identification and employee retention in post-acquisition integration". *Human Relations* (73:7), pp. 981–1009.

Sun, Y., Zhuang, F., Zhu, H., Song, X., He, Q., and Xiong, H. 2019. "The impact of person-organization fit on talent management: A structure-aware convolutional neural network approach". In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1625–1633.

Teng, M., Zhu, H., Liu, C., Zhu, C., and Xiong, H. 2019. "Exploiting the contagious effect for employee turnover prediction". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 01, pp. 1166–1173.

Thanos, I. C. and Papadakis, V. M. 2012. "The use of accounting-based measures in measuring M&A performance: a review of five decades of research". In: *Adv. in Mergers and Acquisitions*. Vol. 10, pp. 103–120.

Thorndike, R. L. 1953. "Who belongs in the family". In: *Psychometrika*. Citeseer.

Trichterborn, A., Zu Knyphausen-Aufseß, D., and Schweizer, L. 2016. "How to improve acquisition performance: The role of a dedicated M&A function, M&A learning process, and M&A capability". *Strategic Management Journal* (37:4), pp. 763–773.

Tuch, C. and O'Sullivan, N. 2007. "The Impact of Acquisitions on Firm Performance: A Review of the Evidence". *International Journal of Management Reviews* (9:2), pp. 141–170.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. 2018. "Graph Attention Networks". In: *Proceedings of the International Conference on Learning Representations*.

Wang, L. and Zajac, E. J. 2007. "Alliance or acquisition? A dyadic perspective on interfirm resource combinations". *Strategic Management Journal* (28:13), pp. 1291–1317.

Westerman, J. W. and Cyr, L. A. 2004. "An integrative analysis of person-organization fit theories". *International Journal of Selection and Assessment* (12:3), pp. 252–261.

Yang, C., Xiao, Y., Zhang, Y., Sun, Y., and Han, J. 2020. "Heterogeneous Network Representation Learning: Survey, Benchmark, Evaluation, and Beyond". *CoRR* (abs/2004.00216). arXiv: 2004.00216.