Multi-Stage Machine Learning Model for Hierarchical Tie Valence Prediction

KARANDEEP SINGH, Data Science Group, Institute for Basic Science
SEUNGEON LEE, Data Science Group, Institute for Basic Science, and School of Computing, KAIST
GIUSEPPE (JOE) LABIANCA, Department of Management, UMass Amherst, and Department of Management, University of Exeter
JESSE MICHAEL FAGAN, Department of Management, University of Exeter
MEEYOUNG CHA, Data Science Group, Institute for Basic Science, and School of Computing, KAIST

Individuals interacting in organizational settings involving varying levels of formal hierarchy naturally form a complex network of social ties having different tie valences (e.g., positive and negative connections). Social ties critically affect employees' satisfaction, behaviors, cognition, and outcomes—yet identifying them solely through survey data is challenging because of the large size of some organizations or the often hidden nature of these ties and their valences. We present a novel deep learning model encompassing NLP and graph neural network techniques that identifies positive and negative ties in a hierarchical network. The proposed model uses human resource attributes as node information and web-logged work conversation data as link information. Our findings suggest that the presence of conversation data improves the tie valence classification by 8.91% compared to employing user attributes alone. This gain came from accurately distinguishing positive ties, particularly for male, non-minority, and older employee groups. We also show a substantial difference in conversation patterns for positive and negative ties with positive ties being associated with more messages exchanged on weekends, and lower use of words related to anger and sadness. These findings have broad implications for facilitating collaboration and managing conflict within organizational and other social networks.

CCS Concepts: • **Applied computing** → **Business Intelligence**; • **Information systems** → **Enterprise information systems;**

Additional Key Words and Phrases: Signed link prediction, sentiment embeddings, graph neural networks, tie-valence prediction, organizational social network

K. Singh and S. Lee contributed equally to this research.
This work was supported by the Institute for Basic Sciences (IBS), Republic of Korea, under IBS-R029-C2 (Singh, Lee, and Cha).
Authors' addresses: K. Singh, Data Science Group, Institute for Basic Science, Daejeon 34126, South Korea; email: karandeepbrar@gmail.com; S. Lee and M. Cha (corresponding author), Data Science Group, Institute for Basic Science, Daejeon 34126, and School of Computing, KAIST, Daejeon 34141, South Korea; emails: archon159@kaist.ac.kr, meeyoungcha@kaist.ac.kr; G. (Joe) Labianca, Department of Management, UMass Amherst, Amherst MA 01003, and Department of Management, University of Exeter, Exeter EX4 4PY, UK; email: joelabianca@gmail.com; J. M. Fagan, Department of Management, University of Exeter, Exeter EX4 4PY, UK; email: aftersox@gmail.com.

1556-4681/2023/02-ART83 $15.00
https://doi.org/10.1145/3579096

ACM Reference format:
Karandeep Singh, Seungeon Lee, Giuseppe (Joe) Labianca, Jesse Michael Fagan, and Meeyoung Cha. 2023. Multi-Stage Machine Learning Model for Hierarchical Tie Valence Prediction. *ACM Trans. Knowl. Discov. Data.* 17, 6, Article 83 (February 2023), 20 pages.
https://doi.org/10.1145/3579096

1 INTRODUCTION

There is growing interest in understanding the role of positive and negative network ties or links—recurring relationships that involve enduring valenced interpersonal judgments—in explaining actors' attitudes, behaviors, cognition, and outcomes [25]. Identification of tie valences can be helpful for a variety of practical downstream tasks such as recommending products [26, 51] or friends [22], estimating the impact of a publication [7], and predicting the dynamics of complex social networks [24, 49]. However, despite the amount of time that individuals spend within social structures involving hierarchy and some level of competition, such as work organizations, little is known about tie valence within these networks. Work contexts can, for example, control which individual can gain a promotion; create competition for scarce organizational resources; place people into specialized units that are often at odds with other units; and introduce power relationships that are difficult to ignore. While viewing organizations as a nexus of social relationships has gained ground in the past decade, the use of large-scale data for tie valence detection in a setting involving hierarchy and some conflict is still in its infancy.

What has made this research and its applications particularly challenging is that negative ties are counter-normative (i.e., are generally frowned upon) and are often hidden from view [37]. While colleagues in the same rank might notice conflicts, these conflicts are intentionally suppressed from higher-ranked individuals. This produces a phenomenon where top management is often unaware of important conflicts occurring below them (e.g., interpersonal or interdepartmental disputes) that might need active attention because these conflicts can grow and draw in others [42], ultimately threatening the organization's proper functioning [3]. Thus, a critical task is to understand where there are positive ties that can encourage collaboration and employee attachment to the organization and negative ties that undermine organizational solidarity and goal achievement. Numerous previous studies have shown that the number of positive and negative ties in an organization affects individual and group-based outcomes, including work performance, job satisfaction, and employee turnover [30].

Electronically-mediated communication and electronic information exchange patterns could be reflective of the tie-valences between the concerned parties. Additionally, the advent of digitization has enabled data collection at an unprecedented scale which could be mined for discovering hidden patterns of interest. This, coupled with the recognition of the importance of understanding workplace ties, has resulted in the development of various data-driven solutions like Microsoft Workplace Analytics [39], OrgMapper [38], and Humanyze Workplace Analytics [28]. These tools utilize corporate data to drive workplace improvements based on a broad set of features, including those from e-mails, meetings, collaboration activities, unscheduled calls, and instant messages. In some cases, these providers already collect these data legally on behalf of the client organizations for other purposes. These digital information exchanges can be used to learn relationship patterns and create actionable insights while still protecting individuals' privacy. Yet much of this potential remains underutilized because tie valence prediction is not incorporated into social network analytic tools.

Given the potential offered by these relatively unexplored data, we propose a computational and data-driven approach to the problem of tie-valence prediction in networks involving hierarchy and competition. We utilize disparate sources of real data from an organization and propose a neural network model that extracts the relational sentiment information from unstructured and structured data sources. More specifically, our model, called **exTV** (Model to **ex**tract **T**ie **V**alences), is trained on anonymized work conversation data, employees' **human resources** (**HR**) data, and a sociometric survey of positive and negative ties among a subsample of members. exTV is the first of its kind to utilize anonymized official conversation texts exchanged between members of varying ranks in an organizational hierarchy to learn the numeric embeddings that are representative of people's sentiments toward one another. This step employs a meticulously designed **natural language processing** (**NLP**) algorithm to handle unstructured textual information and other critical meta information such as messaging frequency and times. Next, these embeddings are used alongside the HR data to build a neural network that identifies relational information among the people and finally delivers the output to be used as the tie valence label.

We explore the model and the underlying data with the results obtained and test what sources and types of data are important for the final classification. In doing so, we discover intricate relationships among individuals of different ranks (e.g., supervisor and subordinate) and individuals of particular profiles (e.g., managers, females, ethnic minorities). We also find that work conversations are a rich resource that uncovers valuable information for discerning additional, and otherwise ambiguous, tie-valences. Our model, exTV, can classify tie valence in the studied organizational network with an AUROC score of **0.8190**. We intend to apply our model in the future to understand the rest of our large organization's network tie valences. Our model can offer key insights into downstream tasks such as improving organizational restructuring and post-merger integration, increasing workplace attachment, decreasing conflict and employee attrition rates, leading to improved organizational functioning and employee well-being.

Ethical considerations. Analyzing interpersonal connections, in general, and tie valence, in particular, can provide valuable insights into the workplace environment even while being conducted ethically in a manner that protects individual privacy. One key is separating the firm gathering the data and conducting the network analyses (the third-party provider) from the client organization that receives recommendations, much as is done in numerous other contexts (e.g., third-party firms often handle client companies' sexual harassment or whistleblower allegations to maintain confidentiality for all parties and provide neutral assessments). The client only receives anonymized, aggregated results that, while actionable, protect individuals' data privacy (e.g., identifying where there is increasing interdepartmental conflict in the organization to initiate conflict management techniques quickly). These data are obtained legally because, in most of the world, the data are owned by companies, and employees are made aware that anything they transmit over company networks can be accessed by those companies (with the European Union being the main notable exception).

2 RELATED WORK

Numerous social theories have been proposed to understand positive and negative ties in networks, including in organizational contexts [50]; these theories are often used when attempting tie valence detection. The most commonly used theory is structural balance theory [8, 14], along with its newer variants such as [44]. Sentiment lexicons like [21, 40] have also been a popular choice for linguistic-features-based sentiment-classification. Other popular methods in link prediction such as Jaccard coefficient, resource allocation index, and preferential attachment can be used to predict "missing" links [35]. In this work, we focus on the tie-valence of existing edges in an organizational social network.

From a computational perspective, work on predicting network tie valence in organizational contexts is nearly non-existent. Though there is plenty of research on link prediction in social networks, transferring this to organizational contexts is impracticable due to different network dynamics in a formal work setting (e.g., one with competing/collaborating departments and a formal reporting hierarchy). From a machine learning point-of-view, GNN [41] based models can operate on and learn representations of graph-structured data and have shown improved performance over traditional deep learning approaches. The strength of GNNs comes from their ability to implicitly learn the graph's structure and the neighboring contextual information. Additionally, works like SGCN [16], SiGNet [29], and SiGAT [27] have adopted GNNs to handle directional and signed networks. These approaches often incorporate social balance theories into the training process, thereby amalgamating computational and sociological approaches. For instance, SGCN includes structural balance theory, and SiGAT captures both the structural balance theory and status theory [15].

Matrix factorization has also been used commonly for analyzing networks and link prediction tasks [2, 4]. For example, [4] proposes a matrix factorization-based model that also cashes in on the users' personality information. Reference [32] rethinks the problem of link prediction by identifying "no-relation" as a possible future status of the node pairs. Studies like [5, 6] introduce advanced graph embedding methods with techniques such as preferential random walks. Research has also been focused on using latent factor models for link prediction tasks, such as [43, 49].

3 PROBLEM SETTING

3.1 Dataset

One of the key factors in determining the performance of a machine learning model is the data being utilized for training and testing purposes. Our research problem is not often attempted because employee data from a work organization are rarely available for research consumption. Even rarer is the data pertaining to the employees' liking and disliking of their fellow colleagues. Understandably, these data are limited and relatively very limited for a computational approach that a machine learning model heavily drives. The various data sources that we utilize to supervise the training and assess the model output are discussed in the following paragraphs[1].

Conversation data consist of official conversation exchanges among employees over two years. While the digital conversations could be in any form, such as e-mails, instant messages, meeting invites, or seminar chat logs, we use e-mails in this study. The digital exchange of information among employees forms an information exchange network. Along with the exchange patterns, we also have at our disposal the anonymized text of the conversations, where the data is stripped of any personally identifiable information.

Each record in this data has multiple features, including message text, ID, timestamp, and the hash digests of the sender and receiver IDs. In total, there are 1,403,303 messages exchanged between 3,404 users. Furthermore, using the exchange patterns, we generate an additional 73 **meta features** from conversation data that may reflect the link polarity, e.g., the number of total e-mails between a user pair, the number of weekend e-mails.

Human resources (HR) data contain the demographic and work-related attributes for each employee. Features in these data are typical of HR data (e.g., age, gender, department, rank, experience). Other information, such as salary, is anonymized due to privacy concerns. The HR data are only available for 1,022 employees (e.g., it does not cover part-time employees or contractors).

Sociometric survey data from the studied organization contain self-reported workplace relationships in a work unit. The survey contained a questionnaire that inquired about the attitudes

[1]Code link: https://github.com/k-s-b/extv.
Due to our non-disclosure agreement with the organization and our Institutional Review Board data management protocol, the raw data cannot be shared.

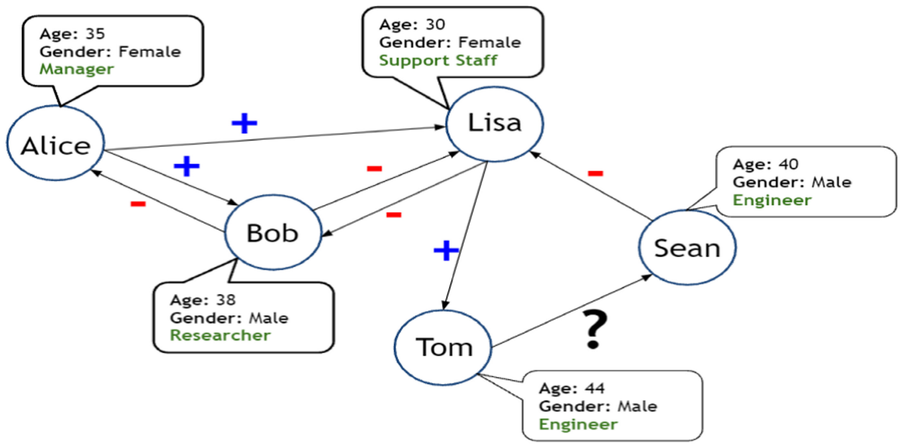


Fig. 1. An example of signed, directed network with partially missing information (i.e., tie valence from Tom to Sean). People with different roles and attributes form the organization depicted.

of **127** employees toward their colleagues. Responses were provided on a seven-point Likert scale, 1 = "dislike a lot" and 7 = "like a lot." We utilize these data to build the dichotomized ground-truth labels for our research problem; specifically, people in an employee's "friend" network have positive labels and the "avoid" network have negative labels. State-of-the-art techniques for valid, reliable, and ethical social network survey data collection, including using optimal question wording, Social Network Data Labs, and institutional ethical oversight were employed [1]. A.3 presents additonal details of the survey.

We reconcile and merge these datasets via employees' anonymized IDs and arrive at the final dataset used as an input to our model. The e-mail and HR data are from an overlapping period, which makes the data unique. This intersection results in data on **98** unique employees with **967** labeled relations among them. Furthermore, these data are split into train, valid, and test sets for model training and analysis purposes. While the final dataset is small in size, our proposed approach is nonetheless able to extract insightful information from the available data annotation. A.1 presents selected exploratory analyses on HR and conversation data.

3.2 Signed Edge Prediction Problem

Employees constantly interact with each other in organizational settings, forming a localized social interaction network. Increasingly, these individuals also exchange digital information (e.g., e-mails, instant messages), forming a virtual social interaction network. The nodes in this network are the individuals, and the traces of digital information exchange would determine edges. When this virtual network is reconciled with the survey responses indicating the perceived link polarity between the people in the offline network, the virtual network's edges could be annotated as positive or negative ties. Overall, this process constitutes a *directed*, *signed*, as well as *node, and edge attributed* social network, where the employees are the nodes and their tie-valences the network's edges. Figure 1 illustrates an example of signed, directed network of our problem.

Given the importance of understanding and predicting the valence of ties between employees, our research problem can be defined as a binary classification task with labels $\{0, 1\}$, denoting negative and positive ties. We formulate this problem with a multi-modal deep-learning approach that uses NLP and graph-based deep-learning techniques. More formally, an organizations' social network can be viewed as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}^+, \mathcal{E}^-)$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ is the set of

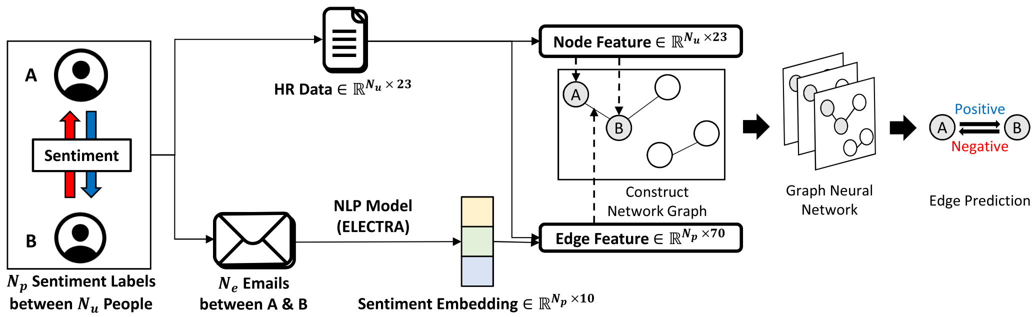


Fig. 2. Overview of the model: exTV consists of two stages—the NLP stage and the semi-supervised signed GNN stage. The NLP stage extracts text embeddings ("Sentiment Embedding" in this figure) with the context of e-mails exchanged between the employee-pairs. The ground truth is the sociometric survey data on the tie-valence between the employees. Network graph is constructed with employees as nodes, and the e-mail exchange as the edges. As the learnt embeddings represent the sentiment between a pair of employees, these embeddings are concatenated with other edge-level feature obtained from the input data. Finally, the node features, updated edge features, and other meta features are leveraged in the semi-supervised signed GNN stage to perform the final classification.

n employees while $\mathcal{E}^+ \subset \mathcal{V} \times \mathcal{V}$ and $\mathcal{E}^- \subset \mathcal{V} \times \mathcal{V}$ represents set of positive and negative edges. A labeled edge between an employee pair, $e_{i,j} : v_i, v_j \in \mathcal{V}$ and $\forall e_{i,j} \in \{\mathcal{E}^+ \cup \mathcal{E}^-\}$, is representative of the sentiment from v_i to v_j. Note that the signs of $e_{i,j}$ and $e_{j,i}$ could be different, as one's perception of a person as a "friend", may not necessarily be reciprocated. $\mathbf{x}_v$ denotes the feature matrix that contains each employee's personal information $\forall v \in V$. $\mathbf{D}^+$ and $\mathbf{D}^-$ are the matrices that contains edge features including conversation information of edges, for $\forall e^+ \in \mathcal{E}^+$ and $\forall e^- \in \mathcal{E}^-$ each.

4 METHODS

The model's first stage processes the raw e-mail text and aims to extract numeric embeddings that represent the underlying reported sentiments. We first explore standard NLP approaches, including sentiment lexicons like LIWC [40] and VADER [21]. Then, we design a neural network that builds upon state-of-the-art deep learning NLP models while accommodating outputs from the above-mentioned standard approaches.

Because the network is composed of sentient people who can influence each other's attitudes, which can, in turn, alter their relationships with other people, we need to add a relational component to the machine learning process. Therefore, in the second stage of exTV, the outputs from stage one, along with other designed meta-features, build upon approaches like graph neural networks and matrix-factorization that take relational information into account while learning the target embeddings. Furthermore, the "signed" models of these approaches also segregate these relational contexts into positive and negative before learning the embeddings to be further utilized for the final classification. The overall model architecture is depicted in Figure 2.

4.1 Text-to-Sentiment Embeddings

This step's goal is to extract numeric embeddings from the e-mail conversation data that are representative of the reported tie-valences among the employees. An employee interacts and exchanges information with other employees via e-mails (and other means), and the nature of this communication is determined by factors such as department affiliation, roles, rank, and perceived tie-valences. We posit that the e-mail text should carry information that is reflective of the nature of the relationship between people and it could be utilized for the final classification task.

We begin by cleaning the conversation data (i.e., e-mail in this study but could be replaced by other types) of unwanted noise by removing signatures and addresses, automated messages, and salutations. The ground-truth labels (the survey data) only exist for a small portion of the otherwise large network with ample conversation text data (approx. 1 million messages in total). Utilizing the whole network and the accompanying exchange data can be potentially advantageous in discovering important hidden features and in enabling the model to "learn" the structure of underlying text.

To enable this utilization, we undertake the fine-tuning of pre-trained and state-of-the-art **ELECTRA** [13] model in an **unsupervised** fashion. This is accomplished by fine-tuning the model with **masked language modeling** (**MLM**). Equation (1) represents the training objective of MLM while Π denotes the index of masked tokens, and X_Π and $X_{-\Pi}$ denotes set of masked tokens and unmasked tokens, respectively [33]

$$L_{mlm}(X_\Pi|X_{-\Pi}) = \frac{1}{K}\sum_{k=1}^{K}\log p(x_{\pi_k}|X_{-\Pi}). \tag{1}$$

As strong sentiments are rarely expressed in workplace messages, the lack of informative signal could potentially lead to the model learning over-smooth sentiment embeddings. Letting our model ingest the entire conversation data customizes the model weights to the mostly neutral, formal conversation style of an organizational workplace.

Thereafter, we employ unsupervised fine-tuned ELECTRA to generate numeric embeddings for all messages pertaining to all pairs of users involved in the e-mail exchange. The e-mail exchange data are naturally unstructured as a variable number of messages of different lengths are exchanged by each person-pair. We design a methodology where this data could be transformed into a fixed-sized embedding for each pair of users in the exchange data.

For each pair, the numeric embeddings for each message obtained from the previous step are stacked and passed through a **multi-head-self-attention** (**MHSA**) layer as described in Equation (2). MHSA step relates different messages in the input and updates the initial embeddings to further "highlight" the underlying sentiment.

Let $\mathbf{F}^Q$, $\mathbf{F}^K$, and $\mathbf{F}^V$ denote matrices for queries, keys, and values in self-attention [45], respectively. *Scaled dot product attention* is defined as

$$SDPA(\mathbf{F}^Q, \mathbf{F}^K, \mathbf{F}^V) = softmax\left(\frac{\mathbf{F}^K(\mathbf{F}^Q)^\top}{\sqrt{d}}\mathbf{F}^V\right), \tag{2}$$

which is used to capture the similarity between the $\mathbf{F}^Q$ and $\mathbf{F}^K$ vectors. Instead of performing a single attention function, the queries, keys, and values can be projected with different learned linear projections, on which the attention mechanism can be performed in parallel. To accomplish this goal of modeling different aspects of interactions between the different messages, multi-head (MH) self-attention is utilized:

$$MH(\mathbf{F}^Q, \mathbf{F}^K, \mathbf{F}^V) = \text{Concat}(head_1, head_2, \ldots, head_{n_h})\mathbf{W}^O, \tag{3}$$

where $head_i = SDPA(\mathbf{F}^Q\mathbf{W}_i^Q, \mathbf{F}^K\mathbf{W}_i^K, \mathbf{F}^V\mathbf{W}_i^V)$, $\mathbf{W}_i^Q, \mathbf{W}_i^K, \mathbf{W}_i^V \in \mathbb{R}^{d\times d_k}$ and $\mathbf{W}^O \in \mathbb{R}^{n_h d_v \times d}$ are learnable weights, and n_h is the number of heads. The dense layers (i.e., $\mathbf{W}_i^Q, \mathbf{W}_i^K, \mathbf{W}_i^V$) are used to project the queries, keys, and values into their vector spaces. Since the queries, keys, and values are all equal to the messages pertaining to a person-pair, i.e., $\mathbf{F}^Q = \mathbf{F}^K = \mathbf{F}^V = \mathbf{S}$, we can produce the multi-head attention-aware sentiment embedding matrix as $\tilde{\mathbf{S}} = MH(\mathbf{S}, \mathbf{S}, \mathbf{S})$.

Input to the MHSA layer is padded (and masked) equal to the length of vector of maximum length in a batch. Next, the output of the MHSA layer is pooled via mean-pooling to arrive at a fixed-sized embedding across all user pairs. Subject to ELECTRA, e-mail messages longer than

ALGORITHM 1: Sentiment Extraction

Input: emails b/w pairs, $\mathbf{m}_{i,j}, \forall i \in \{1 \dots n_j\}, \forall j \in \{1 \dots N\}$, where n_j denotes the number of emails of pair j and N denotes the number of pairs, pre-trained **electra**, set of entire unlabeled emails U

Output: email context embeddings $\mathbf{z}_j^{email}, \forall j \in \{1, \dots, N_p\}$

// Use **LIWC** and **VADER** for aggregated emails per pair.

$\mathbf{F_{LIWC}}(j) := \mathbf{LIWC}(\mathbf{CONCAT}(\mathbf{m}_{i,j}))$;

$\mathbf{F_{VADER}}(j) := \mathbf{VADER}(\mathbf{CONCAT}(\mathbf{m}_{i,j}))$;

// Train a **XGBoost** model, and extract its leaves.

$\mathbf{XL}(f) := \mathbf{LEAVES}(\mathbf{XGB}(f(j))), f \in \{\mathbf{F_{LIWC}}, \mathbf{F_{VADER}}\}, \forall j \in \{1 \dots N\}$

// Unsupervised finetuning of **ELECTRA** via MLM

ELECTRA = MLM(**electra**, U)

for $j \in \{1, 2, \dots, N\}$ **do**

    for $i \in \{1, 2, \dots, n_j\}$ **do**

        // Get email embeddings with **ELECTRA**.

        $e_{i,j} = \mathbf{ELECTRA}(\mathbf{m}_{i,j})$

    end

    // Update **ELECTRA** embeddings by **MHSA**.

    $a_{i,j} = \mathbf{MHSA}(e_{i,j})$

    // Obtain hidden state $\mathbf{h}_j^E$ by pooling from **MHSA** result.

    $\mathbf{h}_j^E = \mathbf{W}^E \mathbf{MEAN}(a_{i,j})$

    // Obtain hidden state $\mathbf{h}_j^L$ from **LIWC**.

    $\mathbf{h}_j^L = \mathbf{W}_0^L [\mathbf{F_{LIWC}}(j), \mathbf{XL}(\mathbf{F_{LIWC}}(j))]$

    // Obtain hidden state $\mathbf{h}_j^V$ from **VADER**.

    $\mathbf{h}_j^V = \mathbf{W}_0^V [\mathbf{F_{VADER}}(j), \mathbf{XL}(\mathbf{F_{VADER}}(j))]$

end

$\mathbf{z}_j^{email} \leftarrow tanh([\mathbf{h}_j^E, \mathbf{h}_j^L, \mathbf{h}_j^V])$

512 words are truncated, and those of shorter lengths are padded, though as shown in Figure 5(b), most e-mail messages are shorter than the upper limit of 512 words and there is rarely any information loss due to this limitation.

To strengthen and aid the sentiment-extraction process, we utilize the outputs of two **sentiment lexicons**: LIWC [40] and VADER [21]. Both LIWC and VADER are fed the concatenated messages for all user-pairs. LIWC reads the input text and outputs the percentage of words that reflect different emotions, thinking styles, and social concerns. VADER considers the polarity and intensity of emotion of the input text and gives four output scores, positive, negative, neutral, and compound (computed by normalizing the other three scores).

We train XGBoost [9] models on the outputs of LIWC and VADER and then let the neural network learn the weights of the concatenated input of original features and the one-hot encoded decision paths in the extracted tree leaves from the trained XGBoost model. This enables the model to learn the relational information between different features of sentiment lexicons.

Finally, ELECTRA embeddings, outputs from sentiment lexicons, and one-hot encoded leaf embeddings from sentiment-lexicon-XGBoost-models are concatenated and fed to an FC layer to produce a fixed-sized embedding vector. The training process is accomplished by performing binary classification against the ground-truth sentiment labels using binary cross-entropy loss.

Multiple experiments established that the larger size of the final embedding vector can lead to overfitting. Hence, we run regularization techniques and maintain a small embedding size to account for the input dataset's size. The sentiment extraction process is presented in Algorithm 1.

4.2 Semi-Supervised Graph Neural Networks

Positive and negative links have different dynamics in a network, and social theories like balance theory offer a systematic way to handle these. Specially-designed machine learning approaches such as the **signed graph neural network models** are the computational counterparts of these sociological approaches, where positive and negative links are initially treated independently, often driven by relevant social theories. The message passing architecture in GNNs in general learns the embedding of a node by leveraging it's own and the aggregated neighboring information:

$$x_i^k = \phi\left(x_i^{k-1}, \gamma(\phi(\mathcal{N}^{k-1}(x_i)))\right), \tag{4}$$

where ϕ is a non-linearity, γ is a permutation invariant function, and $\mathcal{N}^{k-1}(x)$ are the neighboring nodes of the target node in the layer $k-1$. This mechanism is applied in signed networks by segregating the nodes by link polarity, and by contrasting the embeddings of these groups of nodes. Any other associated information such as link direction, topological structure, associated node features can easily be incorporated into the model-training. In this work, we employ the SGCN [16], that incorporates one of the most noteworthy signed social network theory—balance theory [8]—as the base information aggregator. SGCN segregates the positive and negative neighbors based on balance theory and employs a segregated aggregation mechanism as presented in Algorithm 2. Additionally, we explored GNNs like SiGAT [27], that include both the balance theory and another popular signed social network theory—the status theory [19, 48]. We account for neighboring nodes that can be reached via a 3-hop path, a hyperparameter design choice.

We experimented with two different aggregation mechanisms for improving the representational ability of GNNs. First, instead of standard order invariant pooling, we employed an attention-based aggregation mechanism. Second, we designed an aggregator function that can ingest the edge-level features in the node neighborhood. This function enabled the pertinent inclusion of topologically-relevant relational information into the training process. However, we discovered that despite a substantial increase in computational complexity, including both of these approaches in the SGCN didn't lead to a statistically significant performance gain. After extensive experimentation we reached the conclusion that owing to a small-sized dataset, the gains offered by both these approaches are invariably offset by over-fitting.

We train a multi layer perceptron for the final classification by concatenating the sentiment embeddings from the NLP stage, and relational embeddings from the SGCN stage. Algorithm 2 describes the aggregation process.

5 RESULTS

We evaluated the model's performance in various test settings. The merged data were randomly split into training, validation, and test sets (with ratios 0.70, 0.15, and 0.15), resulting in 682, 140, and 145 data points, respectively. As summarized in Tables 1 and 2, model performance for both the NLP stage and for the complete model is compared against strong baselines, including XGBoost, and various signed GNN models. The embeddings from state-of-the-art NLP-only models are used as baselines for establishing the superiority of the sentiment embeddings extracted via the proposed approach. For the complete model, the 10-run averages are reported with and without sentiment-embeddings. We use early stopping on the validation set, and the best Macro-F1, precision, recall, and AUC scores for test sets are reported. For all models, the best model is selected by searching the embedding size in {32,64,128} and number of epochs in {50,100,200}. The results demonstrate our approach outperforms all baselines, achieving the best performance for all reported metrics, along with a best AUROC score of 0.8190. It is worth mentioning here that a logistic regression model (not shown in Tables 1 and 2) yielded a low AUROC score of 0.5610.

ALGORITHM 2: Balance Theory-Based Aggregation

Input: $G = (V, E^+, E^-)$; node feature $\mathbf{x}_i$, $\forall v_i \in V$; neighbor nodes $v_j \in \mathcal{N}_i$, where $\mathcal{N}_i$ denotes the neighbors of v_i; edge feature between positive neighbors $\mathbf{D}^+_{i,j}$, $\forall e_{i,j} \in E^+$; edge feature between negative neighbors $\mathbf{D}^-_{i,j}$, $\forall e_{i,j} \in E^-$; number of layers L; weight matrices $\mathbf{W}^{P(l)}$ and $\mathbf{W}^{N(l)}$, $\forall l \in \{1 \ldots L\}$; activation function σ

Output: node embedding vectors determining tie valence z_i, $\forall v_i \in V$

// (Optional attention based) node aggregation for neighbors

$\mathbf{F}^+_{Att}(l, i) := \sum_{j \in \mathcal{N}^+_i} \mathbf{ATTENTION}(\mathbf{h}_j{}^{P(l)})$

$\mathbf{F}^-_{Att}(l, i) := \sum_{k \in \mathcal{N}^-_i} \mathbf{ATTENTION}(\mathbf{h}_k{}^{P(l)})$

// Initialize the first layer with given node features

$\mathbf{h}_i{}^{(0)} \leftarrow \mathbf{x}_i, \ \forall v_i \in V$

for $v_i \in V$ **do**

  // Aggregate edge information between negative neighbors

  $\mathbf{h}_i{}^{D^+} = \mathbf{POOL}(D^+_{i,j}), \ \forall j \in \mathcal{N}^+_i$

  // Aggregate edge information between negative neighbors

  $\mathbf{h}_i{}^{D^-} = \mathbf{POOL}(D^-_{i,k}), \ \forall k \in \mathcal{N}^-_i$

  // Obtain positive hidden state for the first layer

  $\mathbf{h}_i{}^{P(1)} = \sigma\left(\mathbf{W}^{P(1)}\left[\mathbf{F}^+_{Att}(0, i), \ \mathbf{h}_i{}^{D^+}, \ \mathbf{h}_i{}^{D^-}, \ \mathbf{h}_i{}^{(0)}\right]\right)$

  // Obtain negative hidden state for the first layer

  $\mathbf{h}_i{}^{N(1)} = \sigma\left(\mathbf{W}^{N(1)}\left[\mathbf{F}^-_{Att}(0, i), \ \mathbf{h}_i{}^{D^+}, \ \mathbf{h}_i{}^{D^-}, \ \mathbf{h}_i{}^{(0)}\right]\right)$

end

if $L > 1$ **then**

  for $l \in \{1, 2, \ldots, n_j\}$ **do**

    for $v_i \in V$ **do**

      // Obtain positive hidden state from the previous layer.

      $\mathbf{h}_i{}^{P(l)} = \sigma\left(\mathbf{W}^{P(l)}\left[\mathbf{F}^+_{Att}(l-1, \ i), \mathbf{F}^-_{Att}(l-1, \ i), \quad \mathbf{h}_i{}^{P(l-1)}, \ \mathbf{h}_i{}^{D^+}, \ \mathbf{h}_i{}^{D^-}\right]\right)$

      // Obtain negative hidden state from the previous layer.

      $\mathbf{h}_i{}^{N(l)} = \sigma\left(\mathbf{W}^{N(l)}\left[\mathbf{F}^+_{Att}(l-1, \ i), \mathbf{F}^-_{Att}(l-1, \ i), \quad \mathbf{h}_i{}^{N(l-1)}, \ \mathbf{h}_i{}^{D^+}, \ \mathbf{h}_i{}^{D^-}\right]\right)$

    end

  end

end

$\mathbf{z}_i \leftarrow [\mathbf{h}_i{}^{P(L)}, \mathbf{h}_i{}^{N(L)}], \ \forall v_i \in V$

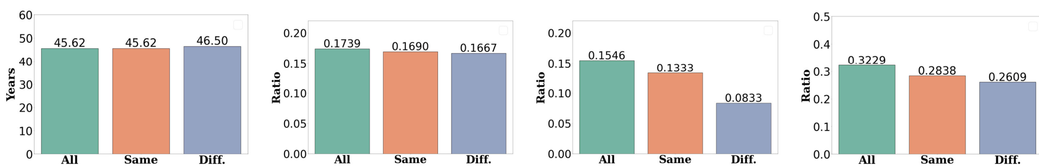
Table 1. Performance Comparison for exTV's Sentiment Embeddings

Model	F1	Precision	Recall	AUC
LIWC & VADER$_{\text{only}}$	0.396	0.328	0.500	0.5
ELECTRA$_{\text{only}}$	0.517	0.576	0.575	0.575
exTV-NLP$_{\text{no_leaves}}$	0.567	0.648	0.579	0.578
exTV-NLP$_{\text{no_meta}}$	0.581	0.601	0.580	0.581
exTV-NLP$_{\text{no_MHSA}}$	0.579	0.618	0.582	0.582
exTV-NLP	0.615	0.650	0.612	0.612

(exTV-NLP). Combinations of the input components also serve the purpose of ablation study.

Table 2. Performance Comparison for with and without Sentiment Embeddings from exTV against Various Baselines

Model	F1	Precision	Recall	AUC
XGBoost	0.625	0.641	0.621	0.621
$\text{XGBoost}_{\text{exTV}}$	0.660	0.668	0.655	0.655
SiGAT	0.586	0.619	0.586	0.666
$\text{SiGAT}_{\text{exTV}}$	0.579	0.581	0.578	0.673
SLF	0.671	0.708	0.662	0.743
SLF_{exTV}	0.676	0.685	0.671	0.765
SGCN	0.680	0.704	0.672	0.781
$\mathbf{SGCN}_{\mathbf{exTV}}$	**0.728**	**0.726**	**0.730**	**0.819**



(a) Median age (b) Ratio of manager (c) Ratio of minority (d) Ratio of females

Fig. 3. Ratios of age, manager, minority, and gender for $\mathcal{S}$ (Same) and $\mathcal{D}$ (Diff.) sets. "All" represents respective values for entire dataset.

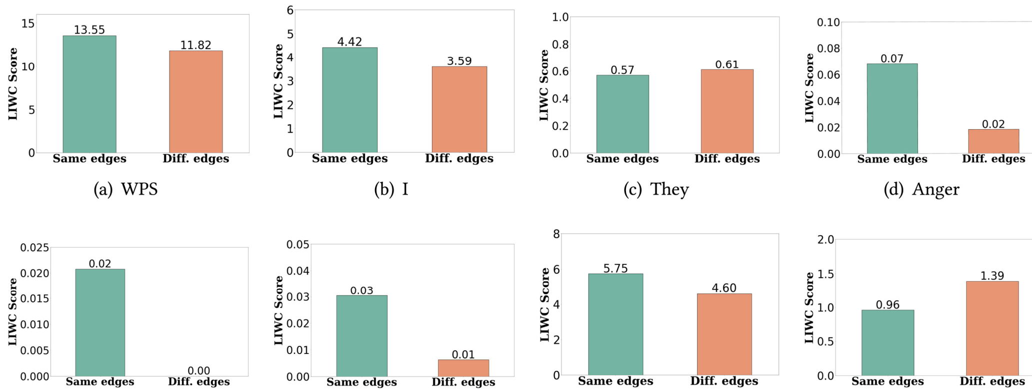
To analyze the functioning of the proposed approach, we further explore model behavior under various test conditions. The inclusion of sentiment embeddings clearly delivers a substantial performance boost over the baselines models. Comparing and contrasting the model output with and without text embeddings will provide an opportunity to discover unique, informative patterns in the underlying e-mail data. In the absence of similar datasets as used in this research, we can not make a direct empirical comparison between such findings from different organizations, but we aim to provide a roadmap for deploying our model to a live setting. The model code is made available publicly.

We chose the best performing model SGCN among the baselines for this analysis. Specifically, the SGCN model was run with and without sentiment embeddings; then **true positive** (**TP**) and **true negative** (**TN**) cases were identified. Furthermore, the TP and TN were compared to achieve a set of data points that were similar (Same or $\mathcal{S}$) or are newly-identified (Diff. or $\mathcal{D}$) in both the runs (with, and without sentiment embeddings). As the performance gain comes from the sentiment information, the new TP and TN can be attributed to new relational patterns uncovered by the e-mail text. Recall that our data-points are the relation-edges between people formed by exchanging digital information. In this light, we perform an exploratory analysis of attributes of relation-edges in the $\mathcal{S}$ and $\mathcal{D}$ sets, as well as the individuals involved in these edges.

5.1 Node-Level Traits

A workplace usually is a unique mixture of individuals from different age groups, nationalities, and ethnicities, who take up different departmental roles according to their knowledge, skills, abilities, and experiences. We explore the HR information of the people in the $\mathcal{S}$ and $\mathcal{D}$ and find that the information uncovered by inclusion of digital exchange data can exhibit certain patterns. Specifically, we analyze age distributions, and manager, minority, and gender ratios.

Figures 3(a)–(d) present the results of this analysis. It can be observed that people in set $\mathcal{D}$ have a higher median age on average. The minority and the female ratio is lower, whereas the ratio of



(a) WPS (b) I (c) They (d) Anger

(e) Sexual (f) Religious (g) Time (h) Money

Fig. 4. Selected LIWC feature values for $\mathcal{S}$ (same edges) and $\mathcal{D}$ (different edges) sets. y-axis of the figure represents LIWC score, which is percentage of words in the text belonging to that dictionary. The values atop bars are the respective percentage values. Each subfigure's title is the feature name in the LIWC output. "WPS" (a) stands for words per sentence, "I" (b) and "They" (c) present the percentage of these pronouns, and similarly, percentage of "Anger" (d), "Sexual" (e), "Religious" (f), "Time" (g), and "Money" (h) words. This figure clearly highlights differences in the nature of communication in the $\mathcal{S}$ and $\mathcal{D}$ sets.

managers is roughly the same. This analysis also points out that in this data, identification of ties for people with relatively higher age, non-minorities, and males have a higher ambiguity, and that utilizing communication data mitigates this ambiguity. Discovery of such patterns can be utilized by management to design better programs and policies that promote better communication among employees, tailored to the organization's needs.

5.2 Edge-Level Features

The edge level features are directly associated with the tie-valence for a pair of individuals, and studying these features can aid the understanding and prediction of the relational edges. We analyze two types of edge features in our data.

5.2.1 Meta Features. We engineered meta-features from the properties of the network formed by the e-mail exchange information. These features were designed with an assumption of being informative of the characteristics of the network formed, indicative of the relationship between a pair of employees, and logically comprehensible. Examples include the total number of e-mails exchanged, average message length, and message frequency. All meta-features are listed in supplementary information Table 3.

Another consideration was to treat the day of week effect. We segregated the exchange for weekdays and weekends to elaborate this pattern further and received intriguing results. For Diff. edges, the average number of e-mails exchanged during weekdays is lower than the Same edges, whereas these numbers are substantially higher during the weekends. Additional exploratory analysis is present in A.1

5.2.2 LIWC Features. LIWC [40] is a non-parametric sentiment lexicon that reveals thoughts, feelings, personality, and motivations based on percentages of the words describing different contexts. LIWC takes as input the text, and outputs a vector of length 93, which are the percentages of words belonging to that dictionary. The individual outputs and their values can be interpreted to analyze the underlying sentiment in the text. We summarize insights from data in Figures 4(a)–(h), that highlight a clear difference in the nature of communication between the $\mathcal{S}$ and $\mathcal{D}$ sets.

The **WPS** (**words per sentence**) feature value is substantially lower for set $\mathcal{D}$, as is the use of first-person singular pronouns (I). However, the use of the third-person plural pronoun (they) is higher. Furthermore, the results also suggest that individuals in the set $\mathcal{D}$ also communicate less about time and more about money. We also find that the use of words concerning anger, sexuality, swearing, religion, with such verbiage being absent from the edges in set $\mathcal{D}$.

5.3 Label-Distributions for $\mathcal{S}$ and $\mathcal{D}$ Sets

We explored the ratio of negative ties for both sets. The values come out to 0.69 and 0.56 for the $\mathcal{S}$ and $\mathcal{D}$ sets, respectively. The newly classified edges exhibit a lower ratio of negative ties or a larger share of positive ties.

The inclusion of e-mail embeddings resulted in better identification of positive edges or emotions. This finding further implies that the final classifier, ingesting the merged information of network and e-mail embeddings, finds an improved signal for the positive relations. It is worth pointing out here that the distribution of labels in the underlying data is imbalanced in favor of negative ties (2:1), potentially resulting in the deep learning model overfitting on this class, mainly due to the smaller data size. Despite this, the analysis reveals that our classifier, and hence the NLP model in stage 1, is finding better text-based indicators for the smaller class. This finding signifies that the text-based positive sentiment is easier to discern than its negative counterpart.

The data show that continual communication on weekends, with a much lower tendency for anger, anxiety, and profanity, signifies a sound positive relation between individuals. Heavy usage of first-person pronouns can be indicative of a preoccupation with one's thoughts and depression [18], and lower usage of such pronouns can suggest a positive relational perspective. Greater use of group connotation pronouns also signals a reduced prevalence of depression [46]. For instance, the usage of "they" is higher in the set $\mathcal{D}$. Similarly, the lower WPS count in Diff. edges might exhibit a more casual form of relationship, akin to friendly acquaintances exchanging information on a digital messaging service, rather than employing long, formal sentences.

The model's superiority stems from sentiment embeddings, as presented in Table 1. Furthermore, detailed comparison of results with and without sentiment embeddings identifies new relationships. Out of the 85 employees in the test data, the inclusion of sentiment embeddings leads to a large performance gain with an accurate tie-valence identification for 16 additional individuals (18.82% of all employees). Such findings can not only greatly aid in shaping the training programs, but also the identification and prediction of the tie-valences in an organization.

6 DISCUSSION AND CONCLUSION

This article presented an ensemble model utilizing anonymized employee information to identify tie-valences in an organizational social network. While ensemble models and their applications have been well researched [10, 11, 17, 20, 31, 34], as per our knowledge, this is the first deep-learning-based ensemble model that leverages archived message text and employee information to learn and predict tie-valences in a context involving hierarchy and organizational structure. While being a data-driven paper, the computational algorithms employed in this work include the most prominent theories in signed social networks—balance theory [8] and status theory [19, 48]. Our model can be applied by third-party providers in a live setting to unobtrusively analyze traffic on an organization's digital communication platforms and deliver useful insights to the client, including providing feedback on emerging interdepartmental conflicts that could threaten the organization's functioning or tracking increased collaboration in a post-merger context [12, 47], all while protecting individuals' anonymity.

This research is built on a snapshot of a much larger dataset and proves that deep learning can be used to better predict and analyze workplace tie valences. Yet, our findings have implications beyond organizational structures and can be used in any online domain, for example, on various online collaboration platforms (e.g., Trello, Slack). Our work suggests insights for future research. As the size of our dataset was relatively small, we expect the model's learning capability to increase greatly when trained over the entire corporate dataset. A larger dataset can be used for modeling. Owing to data limitations, we have employed transductive GNNs for the model's second stage; however, larger datasets will warrant using inductive approaches like GraphSAGE [23]. Similarly, an end-to-end training regimen will allow for improved learning of embedding in both of the stages.

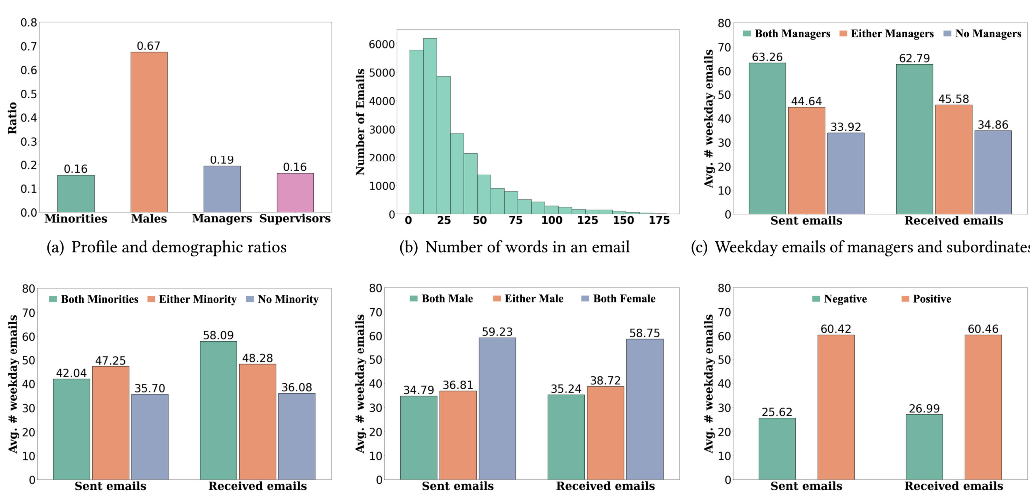
Findings from this work have direct implications in promoting positively valenced relations and collaboration in organizational and other social networks. The proposed method can be employed in any social network by replacing/augmenting the information exchange mechanism, e.g., social media posts, instant messages, bulletin boards, and calendar invites. The relatively high performance of exTV also renders it helpful in pursuing what-if analysis in the social networks.

While the use of personal, conversational data is always fraught with privacy concerns, there are many ways to manage the risk. We suggest separating the providing firm collecting and analyzing the data from the client firm receiving the anonymized, aggregated suggestions for improvement. Combining this with a machine-learning-based method that does not involve humans accessing private data allows the client firm to protect its employees while deriving valuable insights that can improve the organization's functioning and the employees' mental health and career outcomes.

APPENDIX

A APPENDIX

A.1 Exploratory Analysis on HR and Conversation Data



(a) Profile and demographic ratios (b) Number of words in an email (c) Weekday emails of managers and subordinates

(d) Weekday emails of minorities and non-minorities (e) Weekday emails of men and women (f) The label trend of weekday emails

Fig. 5. Exploratory analysis. (a) The ratios of different profiles and demographic features. (b) A histogram of number of words in an e-mail. (c)–(e) Average number of e-mails exchanged between individuals of different profiles and demographic attributes. (f) Average number of e-mails exchanged between edges with different tie-valences (positive and negative). "Out-" and "In-" e-mails denote the e-mails sent and received, respectively.

Figure 5 presents selected exploratory analyses on our dataset. A larger majority of the individuals are males (67%), managers (19%), ethnic minorities (16%), and front-line supervisors (16%). The word count of e-mails has a long-tailed distribution with the peak at 50–100, implying that most of the e-mails involve shorter conversations. Combined with the fact that these exchanges are taking place in a formal setting, this makes mining useful information from text more difficult.

Figures 5(c)–(f) present unique traits—the average number of e-mails exchanged—of the e-mail exchange patterns between different types of user pairs in our data. In Figure 5(c), it can be observed that managers send and receive the most, and subordinates the least information among themselves. This observation could be attributed to the fact that, by the nature of their role, managers interact with many people and tend to participate in many higher-level meetings. Figure 5(d) shows that on average, minorities send out more e-mails to non-minorities (likely driven by numerical probabilities given how few minorities there are), but receive the most e-mails from minorities themselves, suggesting e-mail exchange is also driven by social identities. The results in Figure 5(e) show that women are more likely to exchange with each other. Finally, Figure 5(f) highlights that as compared to negative ties, people with positive workplace relationships interact substantially more with each other. This suggests that reducing negative ties might improve an organization's effectiveness and productivity by promoting better interactions and exchange.

A.2 Meta Features

Table 3. Designed Meta Features

Feature name	Interpretation
total_time_<out/in>	The total period that e-mails have been sent/received.
num_email_<out/in>	The number of e-mails that e-mails have been sent/received.
frequency_<out/in>	The frequency that e-mails have been sent/received. (num_email_out / total_time_out)
avg_interval_<out/in>	The average interval that e-mails have been sent/received. (total_time_out / num_email_out)
med_total_pos_<out/in>	The relative time of median e-mail that have been sent/received in total period.
med_year_pos_<out/in>	The year-wise relative time of median e-mail that have been sent/received.
med_month_pos_<out/in>	The month-wise relative time of median e-mail that have been sent/received.
num_week_emails_<out/in>	The number of e-mails that have been sent/received in weekdays.
num_weekend_emails_<out/in>	The number of e-mails that have been sent/received in weekend.
month_max_num_<out/in>	The maximum number of e-mails that have been sent/received in a month.
month_max_<month>_<out/in>	The maximum number of e-mails that have been sent/received in <month>.
avg_length_<out/in>	The average number of characters in an e-mail that has been sent/received.
avg_sentence_len_<out/in>	The average sentence length in an e-mail that has been sent/received.
avg_sentence_num_<out/in>	The average number of sentences in an e-mail that has been sent/received.

All meta features are computed for sent (out) and received (in) messages. <month> is replaced by all months.

A.3 Survey Description

The ground-truth network used to train the model was collected via survey back in September 2012 in the new product development unit of this consumer product organization's corporate headquarters. Social network analysis was conducted to help reorganize the unit to speed up new product development. A roster with all 185 employees' names was provided and the following social network questions were administered, including the positive tie (friend) and negative tie (avoid) question. Of the 185 sociometric surveys distributed, 144 completed surveys were returned.

- Desired Collaboration
  - I would be more effective in my work if I were able to collaborate more closely with this person. (respondent will check the appropriate names on the roster)
- Friend
  - Do you consider this person to be a close friend (e.g., confide in this person)? (respondent will check the appropriate names on the roster)
- Avoid
  - Sometimes people at work may make us feel uncomfortable or uneasy and, therefore, we try to avoid interacting with them. Do you try to avoid interacting with this person? (respondent will check the appropriate names on the roster)
- Innovation Ratings
  - Innovative employees have the ability to effectively generate and implement novel ideas in the workplace. Please rate how innovative you believe each of your coworkers is. (1 = never innovative, 7 = always innovative)

A.4 Shapley Analysis

We present the SHAP analysis [36] of a XGBoost edge-label classifier that utilizes the same information as the full model. As seen in Figure 6, this analysis could reconfirm that sentiment embeddings are key features in the classification task. It can also be observed that people who send a relatively higher number of messages on weekends are inclined to give positive workplace relationship feedback. Similarly older people tend to give more positive feedback as well. On the other hand, belonging to minority groups seem to be correlated with receiving negative feedback in perceived workplace ties.

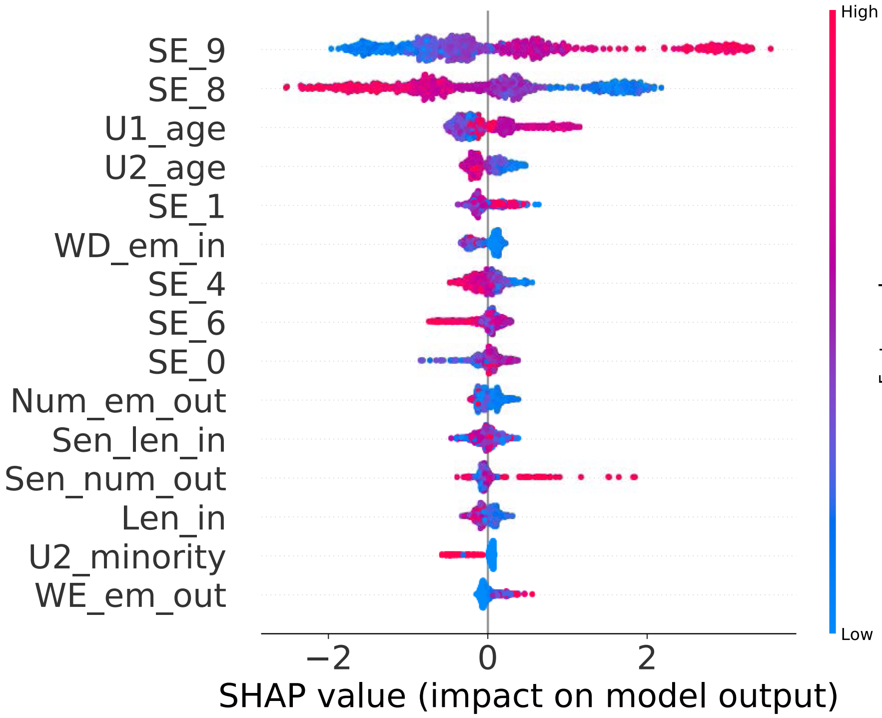
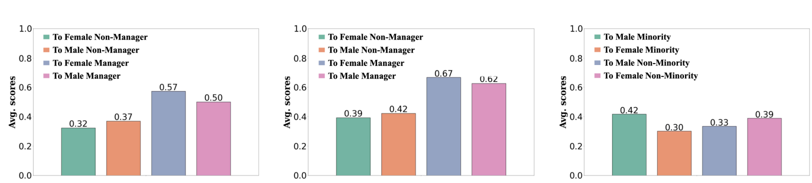


Fig. 6. SHAP summary plot. Left vertical axis - features sorted in descending order of importance. Vertical bar - color legend for feature values. Horizontal axis - how values of different features drive the model output. For display purposes, feature names are shortened. SE_<> : sentiment embeddings, U<>_<age/minority> :user1/user2 age/minority, <WD/WE/Num>_em_<in/out> :number of weekday/weekend/number of e-mails in/out, Sen_<len/num>_<in/out> : sentence length/number in/out.

A.5 Additional Exploratory Analysis

Additional exploratory analysis on the valence-data. Figure 7 presents the mean valence score given by managers. The figure presents segregated scores from male and female managers, as well scores from managers to people in minority and non-minority sets. Figure 8 shows the weekend communication trend for different scenarios. An interesting observation here is the pattern for positive and negative edges, with people with positive edges communicating substantially higher during the weekend “off-hours”.



(a) Male manager (b) Female manager (c) Managers to minorities

Fig. 7. Valence scores by managers.

(a) Managers and subordinates (b) Minorities and non-minorities (c) Gender (d) By label type

Fig. 8. Avg. number of weekend e-mails exchanged.

ACKNOWLEDGMENTS

Authors are grateful for the assistance of the organization's CHRO, as well as all the employees who assisted in this data collection (particularly within the IT department), without whom this research would not have been possible. This work was supported by the Institute for Basic Sciences (IBS), Republic of Korea, under IBS-R029-C2 (Singh, Lee, and Cha), and National Research Foundation of Korea (NRF), under RS-2022-00165347 (Lee and Cha).

REFERENCES

[1] Filip Agneessens and Giuseppe (Joe) Labianca. 2022. Collecting survey-based social network information in work organizations. *Social Networks* 68 (2022), 31–47.

[2] Priyanka Agrawal, Vikas K. Garg, and Ramasuri Narayanam. 2013. Link label prediction in signed social networks. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*. 2591–2597.

[3] Chester Irving Barnard. 1968. *The Functions of the Executive*. Vol. 11. Harvard University Press.

[4] Ghazaleh Beigi, Suhas Ranganath, and Huan Liu. 2019. Signed link prediction with sparse data: The role of personality information. In *Proceedings of the 2019 World Wide Web Conference*. 1270–1278.

[5] Kamal Berahmand, Elahe Nasiri, Saman Forouzandeh, and Yuefeng Li. 2022. A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. *Journal of King Saud University - Computer and Information Sciences* 34, 8, Part A (2022), 5375–5387.

[6] Kamal Berahmand, Elahe Nasiri, Mehrdad Rostami, and Saman Forouzandeh. 2021. A modified DeepWalk method for link prediction in attributed social network. *Computing* 103, 10 (2021), 2227–2249.

[7] Xiaoyan Cai, Junwei Han, and Libin Yang. 2018. Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 5747–5754.

[8] Dorwin Cartwright and Frank Harary. 1956. Structural balance: A generalization of Heider's theory. *Psychological Review* 63, 5 (1956), 277.

[9] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 785–794.

[10] Yi-Ling Chen, Ming-Syan Chen, and Philip S. Yu. 2015. Ensemble of diverse sparsifications for link prediction in large-scale networks. In *Proceedings of the 2015 IEEE International Conference on Data Mining*. 51–60.

[11] Yen-Liang Chen, Chen-Hsin Hsiao, and Chia-Chi Wu. 2022. An ensemble model for link prediction based on graph embedding. *Decision Support Systems* 157 (2022), 113753.

[12] Chia-Yen (Chad) Chiu, Prasad Balkundi, Bradley P. Owens, and Paul E. Tesluk. 2020. Shaping positive and negative ties to improve team effectiveness: The roles of leader humility and team helping norms. *Human Relations* 75, 3 (2020), 502–531.

[13] Kevin Clark, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning. 2020. ELECTRA: Pre-training text encoders as discriminators rather than generators. In *Proceedings of the International Conference on Learning Representations*.

[14] James A. Davis. 1967. Clustering and structural balance in graphs. *Human Relations* 20, 2 (1967), 181–187.

[15] James A. Davis and Samuel Leinhardt. 1970. The structure of positive interpersonal relations in small groups. *Sociological Theories in PRogress* 2 (1970).

[16] Tyler Derr, Yao Ma, and Jiliang Tang. 2018. Signed graph convolutional networks. In *Proceedings of the 2018 IEEE International Conference on Data Mining*. 929–934.

[17] Liang Duan, Shuai Ma, Charu Aggarwal, Tiejun Ma, and Jinpeng Huai. 2017. An ensemble approach to link prediction. *IEEE Transactions on Knowledge and Data Engineering* 29, 11 (2017), 2402–2416.

[18] To'Meisha Edwards and Nicholas S. Holtzman. 2017. A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality* 68 (2017), 63–68.

[19] M. Hamit Fişek, Joseph Berger, and Robert Z. Norman. 1991. Participation in heterogeneous and homogeneous groups: A theoretical integration. *American Journal of Sociology* 97, 1 (1991), 114–142.
[20] Saman Forouzandeh, Kamal Berahmand, Elahe Nasiri, and Mehrdad Rostami. 2021. A hotel recommender system for tourists using the artificial bee colony algorithm and fuzzy TOPSIS model: A case study of TripAdvisor. *International Journal of Information Technology & Decision Making* 20, 01 (2021), 399–429.
[21] CHE Gilbert and Erric Hutto. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*.
[22] Avni Gulati and Magdalini Eirinaki. 2019. With a little help from my friends (and their friends): Influence neighborhoods for social recommendations. In *Proceedings of the World Wide Web Conference*. 2778–2784.
[23] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 1024–1034.
[24] Xiaorong Hao, Tao Lian, and Li Wang. 2020. Dynamic link prediction by integrating node vector evolution and local neighborhood representation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1717–1720.
[25] Nicholas M. Harrigan, Giuseppe (Joe) Labianca, and Filip Agneessens. 2020. Negative ties and signed graphs research: Stimulating research on dissociative forces in social networks. *Social Networks* 60 (2020), 1–10.
[26] Xiangnan He, Hanwang Zhang, Min-Yen Kan, and Tat-Seng Chua. 2016. Fast matrix factorization for online recommendation with implicit feedback. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 549–558.
[27] Junjie Huang, Huawei Shen, Liang Hou, and Xueqi Cheng. 2019. Signed graph attention networks. In *Proceedings of the International Conference on Artificial Neural Networks*. 566–577.
[28] Humanyze. 2011. Humanyze Workplace Analytics. Retrieved from https://humanyze.com/.
[29] Mohammad Raihanul Islam, B. Aditya Prakash, and Naren Ramakrishnan. 2018. SIGNet: Scalable embeddings for signed networks. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*. 157–169.
[30] Giuseppe Joe Labianca. 2014. Negative ties in organizational networks. In *Contemporary Perspectives on Organizational Social Networks*. Emerald Group Publishing Limited.
[31] Kuanyang Li, Lilan Tu, and Lang Chai. 2020. Ensemble-model-based link prediction of complex networks. *Computer Networks* 166, C (2020), 106978.
[32] Xiaoming Li, Hui Fang, and Jie Zhang. 2017. Rethinking the link prediction problem in signed social networks. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. 4955–4956.
[33] Yi Liao, Xin Jiang, and Qun Liu. 2020. Probabilistically masked language model capable of autoregressive generation in arbitrary word order. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 263–274.
[34] Yu lin He, James N. K. Liu, Yan xing Hu, and Xi zhao Wang. 2015. OWA operator based link prediction ensemble for social network. *Expert Systems with Applications* 42, 1 (2015), 21–50.
[35] Linyuan Lü and Tao Zhou. 2011. Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and Its Applications* 390, 6 (2011), 1150–1170.
[36] Scott M. Lundberg, Gabriel G. Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2, 1 (2020), 56–67.
[37] Joshua E. Marineau and Giuseppe Joe Labianca. 2021. Positive and negative tie perceptual accuracy: Pollyanna principle vs. negative asymmetry explanations. *Social Networks* 64 (2021), 83–98.
[38] Maven7. 2017. Maven7 OrgMapper. Retrieved from https://orgmapper.com/maven7/.
[39] Microsoft. 2023. Microsoft Viva. https://www.microsoft.com/ko-kr/microsoft-viva. Accessed January 31, 2023.
[40] James W. Pennebaker, Roger J. Booth, Ryan L. Boyd, and Martha E. Francis. 2015. Linguistic Inquiry and Word Count: LIWC2015.
[41] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. *IEEE Transactions on Neural Networks* 20, 1 (2008), 61–80.
[42] Kenwyn K. Smith. 1989. The movement of conflict in organizations: The joint dynamics of splitting and triangulation. *Administrative Science Quarterly* 34, 1 (1989), 1–20.
[43] Yiyi Tao, Yiling Jia, Nan Wang, and Hongning Wang. 2019. The FacT: Taming latent factor models for explainability with factorization trees. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 295–304.
[44] Wiebe Van der Hoek, Louwe Kuijer, and Yì Wáng. 2020. Logics of allies and enemies: A formal approach to the dynamics of social balance theory. In *Proceedings of the International Joint Conference on Artificial Intelligence*.
[45] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 5998–6008.

[46] Nikhita Vedula and Srinivasan Parthasarathy. 2017. Emotional and linguistic cues of depression from social media. In *Proceedings of the 2017 International Conference on Digital Health*. 127–136.
[47] Vijaya Venkataramani, Giuseppe Joe Labianca, and Travis Grosser. 2013. Positive and negative workplace relationships, social satisfaction, and organizational attachment. *Journal of Applied Psychology* 98, 6 (2013), 1028.
[48] David Willer. 1999. *Network Exchange Theory*. Greenwood Publishing Group.
[49] Pinghua Xu, Wenbin Hu, Jia Wu, and Bo Du. 2019. Link prediction with signed latent factors in signed social networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1046–1054.
[50] Janice Yap and Nicholas Harrigan. 2015. Why does everybody hate me? Balance, status, and homophily: The triumvirate of signed tie formation. *Social Networks* 40 (2015), 103–122.
[51] Muhan Zhang and Yixin Chen. 2020. Inductive matrix completion based on graph neural networks. In *Proceedings of the International Conference on Learning Representations*.

Received 24 March 2022; revised 27 October 2022; accepted 16 December 2022