More than Meets the Tie: Examining the Role of Interpersonal Relationships in Social Networks

Minje Choi, Ceren Budak, Daniel M. Romero, David Jurgens

University of Michigan

{minje,cbudak,drom,jurgens}@umich.edu

Abstract

Topics in conversations depend in part on the type of interpersonal relationship between speakers, such as friendship, kinship, or romance. Identifying these relationships can provide a rich description of how individuals communicate and reveal how relationships influence the way people share information. Using a dataset of more than 9.6M dyads of Twitter users, we show how relationship types influence language use, topic diversity, communication frequencies, and diurnal patterns of conversations. These differences can be used to predict the relationship between two users, with the best predictive model achieving a macro F1 score of 0.70. We also demonstrate how relationship types influence communication dynamics through the task of predicting future retweets. Adding relationships as a feature to a strong baseline model increases the F1 and recall by 1% and 2%. The results of this study suggest relationship types have the potential to provide new insights into how communication and information diffusion occur in social networks.

1 Introduction

Dyadic relationship between individuals are a fundamental characteristic in online social networks such as Twitter. For relationships, concepts such as tie strength (Granovetter 1973), signs (Leskovec, Huttenlocher, and Kleinberg 2010a,b), and direction (Foster et al. 2010) have been used to study communicative behaviors such as reciprocity (Cheng et al. 2011), topic diffusion (Romero, Tan, and Ugander 2013), and echo chambers (Colleoni, Rozza, and Arvidsson 2014). Individuals in these networks are largely organized around social structures such as work, neighborhood or families (Feld 1981, 1982), forming *interpersonal relationships*, such as friendships, kinship, and romantic partnerships. These interpersonal relationship types can influence communication and behavior in the network—e.g., consider what information might be shared between friends versus with a parent. Knowing and inferring relationship types in a social network can have several implications, such as directing messages to the appropriate social audience (Ranganath et al. 2015), improving information diffusion models, and detecting social communities (Tang, Wang, and Liu 2012). However, due to lack of

data availability, interpersonal relationships have rarely been considered for these tasks in social network research.

In this paper, we aim to close this gap by inferring interpersonal relationship types from dyadic interactions in online social networks. Indeed, several studies have tried to classify relationships in domains such as phone call logs (Min et al. 2013), chatroom conversations (Tuulos and Tirri 2004), and conversation transcripts across messaging platforms (Welch et al. 2019). While these studies show predicting relationships is possible for private exchanges or niche topic-based communities, general social media lead to additional challenges due to the substantially higher diversity in content and relationship types. Also, while there has been prior work directly aiming to predict relationships from Twitter data (Adali, Sisenda, and Magdon-Ismail 2012), the predicted categories are data-driven clusters that do not directly correspond to known social relationship types. Yet, as we will show, interpersonal communication still contains linguistic signals that reveal social relationships, enabling accurate prediction.

The contributions of this study are as follows. First, using a massive dataset of interactions between 9.6 million Twitter user dyads with labeled relationships, we conduct an extensive analysis of linguistic, topical, network and diurnal characteristics across relationship categories. We show that relationships on Twitter follow existing theories of interpersonal relationships and reveal complex social dynamics. Second, we introduce a neural network model for classifying five relationship types from linguistic and network features, achieving an F1 of 0.70, that substantially improves upon a strong classifier baseline (0.55) and random guess (0.20). Finally, we show that knowing the type of relationship improves performance on the challenging task of predicting whether one user will retweet another's message, improving the F1 by 1.4% for tweets that do not contain URLs and 2.0% for tweets that do, and highlighting the benefit of modeling the interaction between relationship and content. A pretrained version of our model is publicly available[1].

2 Interpersonal Relationships

Interpersonal relationships between Twitter users can be broadly grouped into five categories: *social*, *romance*, *family*, *organizational* and *parasocial*. These categories, based

---

[1]https://github.com/minjechoi/relationships

Category	Examples
Social	best friend, neighbor, roommate
Romance	dating partner, spouse, fiancé
Family	parent, child, aunt
Organizational	manager, colleague, pastor
Parasocial	idol, fan, hero

Table 1: Examples of relationship types per category

Category	Dyads	DM	PM	RT
Social	6.6M	81M	23.9M	47M
Romance	2.3M	36M	11M	20M
Family	324K	3.4M	945K	1.7M
Organizational	92K	419K	316K	470K
Parasocial	360K	4.1M	3M	4.1M
Total	9.6M	125M	39M	74M

Table 2: Statistics of dyad pairs and tweet interactions in the final dataset for directed mentions (DM), public mentions (PM) and retweets (RT).

on prior theory from communication studies and sociology, cover the social relationships studied in both offline (Knapp, Ellis, and Williams 1980; Feld 1982) and online settings (Ozenc and Farnham 2011). Examples for each category are included in Table 1.

Social Peer relationships and friendships are often the most common relationship in one's social network (Gorrese and Ruggieri 2012). Characteristics include high levels of reciprocity (Hartup and Stevens 1999), a wide range of shared topics (Hays 1984) and homophily (Rivas 2009). Strong ties include close friends who provide emotional support (Richey and Richey 1980), while weaker ties such as acquaintances or neighbors can help build connections and obtain information (Granovetter 1983). Several studies on online social networks have focused on the interactions of social relationships (Ellison et al. 2013; Lee 2009; Burke and Kraut 2014).

Romance Romantic relationships are central to adult life, leading to opportunities for intimacy and support (Hartup et al. 1999). They exist in various stages such as dating, engaged and being married (Stafford and Canary 1991; Knapp, Ellis, and Williams 1980), and can develop into the formation of new families. These relationships are often considered the closest ties. Previous work introduced methods to classify romantic relationships in online social networks based on their network properties (Backstrom and Kleinberg 2014) and conversation content (Tay, Tuan, and Hui 2018).

Family Family relationships are essential for building personalities and receiving social support. Though maintained throughout lifetime, their importance may decline and are partially replaced by social and romantic ties during young adulthood (Shulman 1975; David-Barrett et al. 2016). This is reflected in contact frequency, diversity of activities and influence strength, which are lower than romantic relationships and similar to friend relationships (Berscheid, Snyder, and Omoto 1989). Topics shared between family relationships in online social networks typically include advice giving and household issues (Burke, Adamic, and Marciniak 2013).

Organizational Relationships are also formed as individuals join organizations and are assigned roles within them (Sluss and Ashforth 2007; Marwell and Hage 1970). Organizational relationships are a mixture of personal and role relationships (Bridge and Baxter 1992). This dual status leads to a stronger notion of a community or group identity (Klein and D'Aunno 1986) and a lesser sense of trust and

solidarity compared to friend relationships (Myers and Johnson 2004). Information exchange and politeness are expected in conversations (Sias et al. 2012).

Parasocial The final relationship category is highly asymmetrical, consisting of celebrity-fan relationships (Garimella, Cohen, and Weber 2017), involving high levels of affection from one side, resembling friendship or romantic relationships (Kehrberg 2015). Parasocial relationships are especially important to study in social networks such as Twitter, as influential figures with millions of followers can influence which topics go "viral" (Suh et al. 2010; Stever and Lawson 2013).

3 Extracting Relationships

In order to construct a ground truth set of dyads with labeled relationships, we use self-reported relationships between users where a user declared their relationships to another user in a tweet. A similar strategy has been used for extracting social roles from tweets (Bergsma and Van Durme 2013; Beller et al. 2014). We now describe the full procedure.

We begin with a 10% sample of all tweets posted between 2012 and 2019 and remove all non-English tweets using `pycld2`. We search for all instances of the phrase 'my **REL** @username', with **REL** being any string of up to three words. Through this search, we capture public relationship declarations such as "My dear husband @username...". Since this will also capture many phrases that do not correspond to a relationship, our next goal will be to filter out such instances.

First, all phrases occurring <1,000 times in the dataset are removed, as we observed that most of the low frequency terms do not correspond to relationships. This process leaves 1,298 phrases that potentially map to a specific relationship. Next, every phrase was assigned to one of the five relationship categories or labeled as invalid through a two-step annotation process by the authors. For each phrase, annotators were shown 50 phrases and asked to choose which categories (up to two) the phrase belonged to, if any. To aid the annotation task, for each phrase, annotators were given five relationship-signaling tweets that used that phrase. Inter-annotator agreement score was measured by averaging the pairwise Fuzzy Kappa score (Kirilenko and Stepchenkova 2016), which allows for multiple categories selected per item. Annotators obtained a κ=0.69, indicating high agreement. Given the high agreement, the remaining phrases were equally distributed across annotators without overlap.
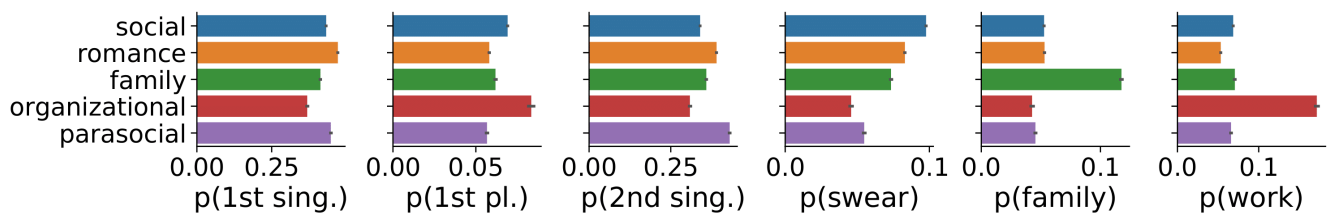
Figure 1: Probability of containing a LIWC-category word in a directed mention to a specific relationship type. Romance and parasocial relationships express high levels of self-disclosure by using more singular pronouns, while organizational relationships use more plural pronouns to show collective identity. Swearing is most common among social and least common within organizational relationships, possibly due to differences in social distance. Work- and family-related words are associated with the respective relationship categories. Here and throughout the paper, error bars denote bootstrapped 95% confidence intervals.

After annotation, phrases assigned either to zero or more than one relationship category were discarded (see Supplemental Material 1 for details)[2]. Ultimately, 508 phrases were assigned to a single relationship category. These phrases were used to label 9,672,541 relationships between 10,410,262 users. Tweets were then collected for all 10.4M users from our 10% sample from 2012–2019, totalling 238M tweets.

Three types of tweets are used in our analyses, which represent different types of interactions between users. (1) **Directed mentions**: Tweets where a user directs a message to a specific user by adding the username at the beginning of the tweet, typically for starting a conversation or replying to another user. While the mentioned user is notified that they were mentioned, this tweet does not appear on the posting user's timeline. (2) **Public mentions**: Tweets visible for the public audience where a username is mentioned in the middle of a tweet. Public mentions are typically used to refer to other users, but not necessarily to have a conversation. (3) **Retweets**: Instances where a user is broadcasting a tweet posted by another user. The number of dyads, directed mentions, public mentions, and retweets for each relationship category are shown in Table 2.

4 Behavioral and Structural Differences in Relationships

To test the quality of extracted relationships, we test communicative and network patterns in each type, validating our data using predictions from known trends for specific relationships (Burke, Adamic, and Marciniak 2013; Ellison, Steinfield, and Lampe 2007).

4.1 Linguistic Preferences

Linguistic style and content reflect how an individual perceives another (Bell 1984). For instance, usage of pronouns reveal levels of self-disclosure (Choudhury and De 2014; Wang, Burke, and Kraut 2016), and swearing terms indicate a closer social distance between the speakers (Feldman et al. 2017). Comparing the use of these words by relationship category can reveal how open each relationship types are

in Twitter conversations. Using lexicons from LIWC (Pennebaker et al. 2015), we calculate the probability of a directed mention containing one of a specific set of words: 1st person singular and plural pronouns, 2nd person singular pronouns, and swearing terms. Assuming that there exist topics central to a single type of relationship such as work-related topics, we also include the LIWC categories for work- and family-related words. The results are displayed in Figure 1.

Communication patterns match prior expectations, as illustrated in three trends. First, conversations in organizational relationships focus on collective identity, as shown in the highest probability of 1st-personal plurals, but lowest in 1st-person singular, as individuals in these relationships associate each other in the context of a larger collective entity (Klein and D'Aunno 1986). Second, parasocial relationships use lesser 1st-person plural pronouns but more 1st- and 2nd-person singular pronouns, making their behavior similar to that of romantic relationships. This result is consistent with previous findings on behaviors in parasocial relationships that resemble love and affection due to the intense focus on the higher status individual (Tukachinsky 2010). However, parasocial conversations also contain substantially fewer swear words compared to romantic relationships, reflecting higher perceived cost of social norm violation due to the relationships larger social distance (Fägersten 2012). Also, consistent with previous findings stating the positive relationship between profanity and social distance (Feldman et al. 2017), swear words appear most commonly for social relationships, followed by romance and family relationships. Finally, the figure also reveals that work- and family-related words match folk expectations: organizational and family relationships are the most likely to use their respective LIWC categories, underscoring the topical differences between relationships. The most frequent five words for each LIWC category (shown in Table 2 in Supplemental Material) confirm that these topical-relationship interactions are not primarily driven by a single word in a category.

4.2 Topical Diversity

Do some relationships talk about more diverse topics than others? Social penetration theory predicts that the variety of topics shared in conversations should increase as relationships further develop (Altman and Taylor 1973). We test this

---

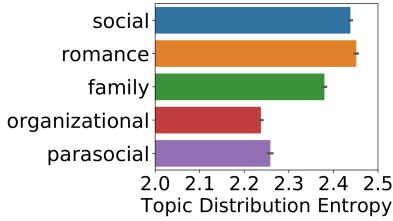[2]Please refer to http://dx.doi.org/10.7302/811 for the supplemental material

Figure 2: The average entropy of topic distributions obtained from directed mention tweets. The entropy is significantly higher for social and romance relationships, which shows these relationships contain more topics in their conversations.
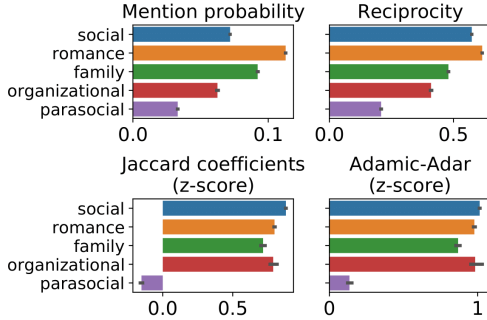

Figure 3: Network and communication features. Jaccard and Adamic-Adar scores are lowest for parasocial relationships, indicating a low similarity in neighbors of a dyad. Romance has both the highest mention probability and reciprocity, signalling the strongest level of mutual communication.

prediction using a topic model to analyze the diversity in communication. Following prior work (Quercia, Capra, and Crowcroft 2012), we measure topical diversity using the entropy of the message's distribution of topics from a trained LDA model. A 100-topic LDA model is fit using Mallet (Mc-Callum 2002) on a sample of 100K dyads balanced across the five categories and using five tweets per dyad to control for differences in communication frequency. To measure diversity we calculate entropy over the mean topic distribution per dyad, then aggregate by category.

The observed topical diversity (Figure 2) matches predictions from social penetration theory, with more diverse topics (higher entropy) seen in relationship categories that are more likely to contain deeper relationships with stronger ties and to have developed further such as romance, social, and family. In contrast, organizational relationships are less likely to communicate on topics outside their common ground (Marwell and Hage 1970). These results were consistent over several runs with topic models trained on 20 or 50 topics.

4.3 Network and Communication Properties

Given their different functions, relationships are expected to differ in their network and communication proprieties. We test for these differences by examining the labeled dyads within the larger social network constructed from our entire

10% Twitter sample. Here, two users have an edge if they both mention each other at least once, which results in a network with ∼1.1B edges. The dyads with labeled relationships represent a small fraction of this comprehensive social network, as the majority of dyads have not declared a relationship. All the network properties for the dyads in our study are measured according to their users' statistics in this larger network. Using this network and directed tweets between two people, we consider two aspects of a relationship: (1) communication frequencies and (2) the local network structure around a relationship.

A comparison of communication frequencies across relationships Communication frequencies are measured using (a) the probability of a user tweeting to the other in a relationship, relative to all others in their ego network, and (b) the reciprocity in communication, measured as the ratio of tweets between two users, scaled to [0,1] where 0 indicates only one person tweets another and 1 is equal communication. We denote $\Gamma(u)$ as the set of neighbors of user u, and $\mathbf{m}_{u \rightarrow w}$ as the number of times user u mentions another user w. The probability of mentioning a specific user out of all possible neighbors is obtained as

$$\frac{\mathbf{m}_{u \rightarrow v}}{\sum_{w \in \Gamma(u)} \mathbf{m}_{u \rightarrow w}}.$$

We also compute the reciprocity between two users as the fraction of communications each user has made, denoted as

$$2 \times \frac{\min \left(\mathbf{m}_{u \rightarrow v}, \mathbf{m}_{v \rightarrow u} \right)}{\mathbf{m}_{u \rightarrow v} + \mathbf{m}_{v \rightarrow u}}.$$

A score of 1.0 means a fully reciprocal dyad with both users communicating equally, and 0 a fully imbalanced dyad where only one mentions the other. To ensure that a relationship is valid, we only calculate reciprocity using dyads where each user has made at least one interaction with the other.

Communication frequencies, which are presented in the first row of Figure 3, exhibit clear differences across categories. Individuals prioritize communication within romantic relationships, consistent with prior work (Burton-Chellew and Dunbar 2015), and have the highest reciprocity. High reciprocity implies two people have similar social status (Verbrugge 1983); this behavior is seen most in categories likely to be between peers: romantic and social relationships. Reciprocity levels follow expectations for differences in social status within each category, with the highest-distance parasocial relationship having lowest reciprocity. Reciprocity relates to status difference, where high reciprocity in contact frequency implies two users belonging to similar social statuses (Verbrugge 1983). While relationships of social and romance categories are genuinely considered as equal in status, other categories contain non-reciprocal relationships such as parent-child relationships or manager-subordinate relationships, which can explain the lower scores for family and organizational categories. Reciprocity is lowest again in parasocial relationships, showing there exist large status differences between celebrity-fan relationships.

By observing mention probability or the likelihood to get mentioned instead of all neighboring users, we can see that

(a) Raw frequency, category-wise

(b) Normalized, category-wise

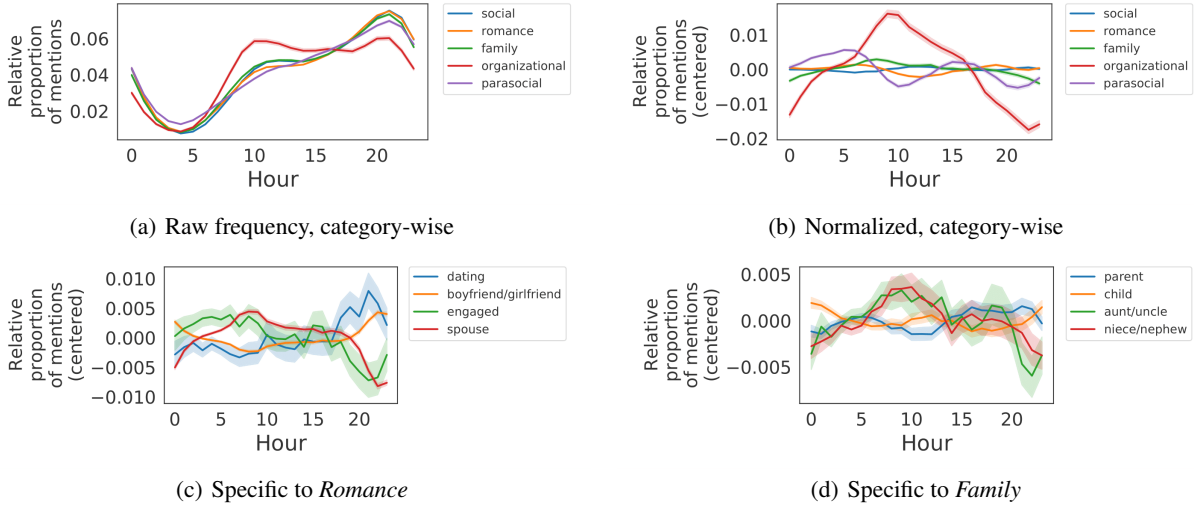(c) Specific to *Romance*

(d) Specific to *Family*

Figure 4: A comparison of mention frequency across hours of day reveal striking difference in temporal dynamics between relationship categories (a,b) and subcategories (c,d) where (b), (c) and (d) are centered relative to the mean temporal distribution across all relationship categories: (a) The un-centered communication frequency among categories (b) the centered communication frequency among categories (c) The centered communication frequency for four *Romance* subcategories (d) The centered communication frequency for four *Family* subcategories. Shaded regions show 95% bootstrapped confidence intervals.

romance has the highest level of relative importance by far, while, surprisingly, social relationships drop to the level of other types. This result is consistent with findings showing that individuals prioritize communicating with their romantic partners over other relationships (Roberts and Dunbar 2011).

A comparison of network properties across relationships We also consider two types of network properties: (a) the Jaccard Index of the two users' friends and (b) Adamic-Adar index (Adamic and Adar 2003), both frequently used for measuring the likelihood of an edge between two nodes. To allow for direct comparisons among dyads, we use the z-normalized score for each metric instead of the raw score as $zscore(u) = \frac{x-\mu}{\sigma}$, where x is the raw score, μ and σ are the mean and standard deviation computed from the neighboring dyads of u other than v. Supplemental Material §3 provides a longer explanation of how these values are computed.

The patterns in network structure (Figure 3 bottom) also match expectations. First, parasocial relationships exhibit very low Jaccard index and Adar-Adamic scores. This is expected as celebrities are embedded in very different social structures from that of their fans and do not have many connections in common. Second, the *family* category has significantly lower Jaccard index and Adar-Adamic score than the social, romance, and organizational categories. This is likely due to two reasons: First, unlike social, romance, and organizational relationships, family relationships do not depend on network structure. Indeed, a family relationship is established regardless of social proximity. However, social, romance, and organizational highly depend on social proximity as these relationships are established through social mechanisms such as friends introducing their friends to each other (i.e. triadic closure) (Kossinets and Watts 2006). Second, family ties tend to be well embedded within fam-

ily networks (e.g. siblings may have other common family connections), but also tend to be much smaller than other relationships such as social and organizational. Due to these differences in volume, family ties are overall less embedded.

4.4 Diurnal Communication Patterns

Individuals manage their communications differently according to social relationships or "identities" such as *social*, *work* and *family* (Ozenc and Farnham 2011; Min et al. 2013). By observing diurnal Twitter usage patterns through the timestamps of tweet messages (Golder and Macy 2011), we show the existence of both between- and within-category communication differences across relationship types.

Using our massive volume of communications, we compute a distribution representing the fraction of messages exchanged for each relationship dyad during each hour of a day. With the number of mentions from user u to user v as $\mathbf{m}_{u \to w}$, we bin each mentioning tweet according to the hour of the day it was created. We then define $t_{u \to w}(i)$ as the fraction of mentions produced in the i-th hour so that $\sum_{i=0}^{23} t_{u \to w}(i) = 1$. We restrict our analysis to tweets where a local timezone of the tweeting user is provided along with its global timestamp and convert the tweet to its local time. Also, we only consider cases where the sending user has made at least 5 activities for better smoothing of the distribution.

After we compute the diurnal communication distributions for each dyad, we can aggregate across different relationship categories or subcategories to obtain category-wise diurnal distributions. We provide a comparison of the diurnal distributions aggregated across different categories, shown in Figure 4(a). While all categories share the same pattern of substantially lower communication during dusk and peaks around evening, similar to previous work (Golder and Macy

2011), we can observe slight differences around daytime. To further examine such differences, we center the distributions by subtracting each with a global mean

$$t_{global}\left(i\right) = \frac{1}{|S|} \sum_{j=1}^{|S|} t_{(u \to v)_j}\left(i\right), \left\{ (u \to v)_j \in S \right\}$$

with S as the set of all dyads across all categories, or in Figures 4(c) and (d), the different subcategories we consider. The centered result (Figure 4(b)) shows a notably higher communication rate for the *Organization* category during work hours (9-16) which drops afterwards, possibly due to moving away from a work activities and towards friends and family chatter (Farnham and Churchill 2011).

Due to our data scale, we can examine communication patterns *within* the same relationship category. The trends for the *Romance* category, shown in Figure 4(c), reveal that even within romance relationships, diurnal patterns can be partitioned into early ("dating" and "boyfriend/girlfriend") versus stable ("engaged" and "spouse") stages, where the former communicates more during late hours. One possible reason is that the latter group may consist of more married couples that share the same physical space during the evening and have fewer reasons to communicate through Twitter. Within family relationships (Figure 4(d)), we show that aunt/uncle - niece/nephew communication is more intense during the day compared to parent-child communication. We conjecture that this tendency reflects the lesser degree of perceived closeness between extended families as opposed to direct kin, restricting communication during late hours.

5 Relationship Classification

Given that relationship types differ in linguistic preferences, topic diversity, and network properties, we now test whether they can be accurately classified from these features.

5.1 Task and Experimental Setup

We classify a dyad into one of the five relationship categories on the basis of its behavioral and communication features. The prediction task is conducted with both balanced and imbalanced datasets. **Balanced set:** Training data uses 200K dyads per category, randomly upsampling organizational relationships which had fewer samples. A random sample of 2K dyads were used for validation data. The test set contained 17,522 dyads per category, where all classes were downsampled to match the least common class, organizational relationships. **Imbalanced set:** 2M dyads are randomly selected and split by a 8:1:1 ratio into training, testing and validation partitions. In both settings, users were constrained to be in only one partition. We ensure that at least one user of the dyad performs at least five interactions, in order to sufficiently represent a dyad's interactions. To control for differences in communication frequency, we restrict the data to at most 15 tweets from each user in a dyad, keeping up to 5 tweets for each type of communication: directed tweets, retweets, and public mentions. To avoid data leakage, all tweets that contained the phrase used for labeling the relationship type were removed prior to this process.

	Model	Soc.	Rom.	Fam.	Org.	Para.	F1
Bal.	*Random*	0.20	0.20	0.20	0.20	0.20	0.20
	GBT Model	0.45	0.57	0.55	0.64	0.55	0.55
	Our Model	**0.60**	**0.69**	**0.69**	**0.79**	**0.72**	**0.70**
Imbal.	*Random*	0.62	0.30	0.05	0.02	0.10	0.20
	Maj. Class	0.76	0.00	0.00	0.00	0.00	0.15
	GBT Model	0.80	0.52	0.33	0.28	0.28	0.44
	Our model	**0.84**	**0.68**	**0.51**	**0.50**	**0.38**	**0.58**

Table 3: Performance comparison for different settings. The F1 score is used to measure the performance in all cases. The first five columns show the F1 scores measured only from samples whose ground truth label belongs to each category as a binary classification task. The last column is the combined Macro F1 score computed as a multi-task classification task.

5.2 Proposed Models

To capture information from the different types of communication, we introduce a new deep learning model that performs multi-level encoding to represent these texts. Our base architecture builds upon the model of Huang and Carley (2019) using the parameters of the RoBERTa pre-trained language model (Liu et al. 2019). Each tweet and bio text for a dyad are encoded as constant-length vectors by encoding each with RoBERTa and mean pooling the output layer's word piece embeddings for the content. These pooled content encodings are fed into a separate 6-layer Transformer network (Vaswani et al. 2017); this second level of encoding summarizes the different sources of text information through its attention mechanism. The output layer of this second-level model is mean pooled as a representation of all communication.

The final dyad representation concatenates the multi-level encoding with (i) character-level embeddings for the username created using 1-dimensional convolutional filters (Kim 2014) and (ii) the four network statistics described in Section 4. This representation is fed through two linear layers using ReLU activation with dropout before a softmax is applied to classify the dyad. RoBERTa models were first pre-trained using 3M training set tweets; then the full classification model was trained end-to-end. Supplemental Material 4.1 provides details of all hyperparameters and training procedures.

Baselines We introduce two baselines: the first is a random guess for all samples, and the stronger second baseline combines text and network features and uses them to train an XGBoost (Chen and Guestrin 2016) classifier. Uni-, bi-, and tri-grams with more than 10K frequency are used as features. We also add features for frequencies in category of the LIWC and Empath (Fast, Chen, and Bernstein 2016) lexicons. Network features are identical to the neural model. Details are included in Supplemental Material §4.2.

5.3 Results and Findings

Our model can accurately recognize different relationships, attaining a macro F1 of 0.70 on the balanced dataset (Table 3), indicating that relationship categories are identifiable from their network and communication patterns. This performance substantially improves upon that of the XGBoost baseline and random chance. In the balanced setting, the model is

most accurate at predicting organizational and parasocial and least at social relationships. We attribute this difference to the intra-class diversity; while organizational and parasocial relationships typically have narrowly-exhibited behavior (e.g., low topic diversity for organizational in Figure 2), social relationships can take many forms, e.g., friends, neighbors. This diversity likely makes the class harder to distinguish.

In the imbalanced setting that reflects the natural distribution of classes, our model offers an even larger performance improvement over baselines. Here, most dyads have a social relationship (76%; cf. Table 3), yet the model is still able to reliably identify all classes. This result highlights the applicability of the model in real-world settings, which is crucial for studying communication dynamics.

Separate ablation studies were performed to test the information from each type of communication (direct, public mention, or retweet) and for the addition of user and network features. Among communication types, public mentions provided the most information (highest performance) for predicting relationship types (0.56 in balanced setting). We speculate that a user has to include context about the mentioned user and their relationship when mentioning them in a public tweet, to provide an explanation to the audience. This information is not required in directed mentions since the expected audience is only the two users. The addition of user features and network features to text features also significantly improves performance, most notably for parasocial relationships, where adding user profiles and network features boosts the F1 score by 0.13. The significantly lower Jaccard coefficients and Adamic-Adar scores of dyads in the parasocial category (Figure 3) likely makes it easier to identify the relationship type, when incorporated as features. Table with full results is in Supplemental Material (Table 3).

5.4 Testing the Validity of Classifier Models

As a further validation of the relationship classification model, we test whether the relationships inferred by our model mirror the behavioral properties seen in the labeled relationships. A random sample of 1M dyads is collected where one user has made at least five interactions, mirroring our classifier setup (§5.1). We then collect mention, reply, and retweet activities made between both users of the dyad, and apply the classifiers on these new users. As temporal information is not used in the classifier but shows clear differences by relationship (Figure 4), we test whether the communication patterns in these inferred relationships are similar. The resulting time series from inferred users were highly correlated with those of the labeled data correlations, ranging from 0.928 to 0.947, shown in detail in Supplemental Table 5 and visualized in Supplemental Figure 1. This result indicates the model infers relationships that have highly similar behavior to the labeled data in practice (despite not being trained on these features), e.g., dyads in the organizational category focusing more of their communication during daytime and dropping in volume after work hours for both labeled and inferred data.

As a second test of validity, we examine the distribution of inferred relationships in the random sample. The resulting distribution differs substantially from that of the labeled data, as expected: Social 23%, Romance 23%, Family 14%,

Organizational 5%, and Parasocial 36% (cf. Table 2). In the random sample we observe more Parasocial, which aligns with earlier expectations that Twitter is largely a mass media platform rather than a social network (Kwak et al. 2010). However, unlike expectations from earlier work, we observe a significant uptick in stronger social relationships: Social and Romance together account for ~46% of the random sample. We attribute this result, in part, to the requirement that dyads in the random sample must have at least five directed tweets, which likely increases the presence of stronger ties who are more likely to talk more. Our results point to the social nature of Twitter and the need for future work to examine how all relationships are manifested on Twitter—not just those that communicate—to establish to what degree the platform now serves as a social network.

6 Retweet Prediction with Relationships

Retweeting is central to information spread on Twitter. Given its significance, several studies have introduced approaches to model and predict this behavior. Factors identified include: content-based features (e.g., hashtags, URLs (Bakshy et al. 2011)), network features (e.g., tie strength (Yuan et al. 2016)), and user popularity (Hong, Dan, and Davison 2011). Despite all these efforts, retweet prediction remains a notoriously hard problem (Martin et al. 2016). Our findings in Section 5 suggest one potential way to improve past efforts. The fact that relationships can be predicted based on tweet conversations indicates a promising connection between the topic of a tweet and the type of relationship that would have interest in it. This leads us to hypothesize that a user's probability of retweeting depends on the interaction between the tweet's content and the relationship to the tweet's author. As such, we test whether incorporating the relationship type between two users u and v can improve the prediction accuracy for whether user v will retweet a particular tweet by u.

6.1 Dataset

We first select the same number of dyads per relationship category to balance across different relationship types. For each dyad we collect all retweets that occurred between the two users, but remove instances where user mentions occurred in the original tweet. While being mentioned in a tweet is a strong motivator for a retweet to occur (Jenders, Kasneci, and Naumann 2013), here our goal is to understand the interplay between the content of the tweet and the relationship properties of the dyad and thus remove tweets with mentions.

We focus on a balanced prediction task where for each positive retweet that occurred between a dyad, we assign one tweet produced by the same user around the same time which did not get retweeted by the other user. As a result, 50,000 positive and negative tweets were sampled per category and split into training, validation, and test sets at a ratio of 8:1:1.

6.2 Models for Retweet Prediction

The models used for the prediction task are also based on a pretrained RoBERTa model for text classification.

Model	Soc.	Rom.	Fam.	Org.	Para.
Baseline	0.61	0.63	0.60	0.64	0.61
With relationships	0.64	0.65	0.63	0.64	0.62

Table 4: Classwise F1 performance comparison of retweet prediction task on a balanced dataset. The presented order is Social, Romance, Family, Organizational, and Parasocial.

	without URLs			with URLs		
Model	Pre.	Rec.	F1	Pre.	Rec.	F1
Baseline	0.58	0.69	0.63	0.53	0.85	0.65
With relationships	0.58	0.71	0.64	0.53	0.87	0.66

Table 5: Model performance (Precision, Recall, F1) at predicting retweets of messages with or without URLs. The addition of relationship types leads to an increase in the F1 score by boosting recall.

Baseline model encodes the tweet text using RoBERTa and preserves the embedding of the first position corresponding to the [CLS] token, which is a common practice in BERT-based classification tasks (Devlin et al. 2018). Features for the baseline model are the contents of the tweet and the number of followers (log-scaled) and the existence of a URL, both frequently used features in retweet prediction (Petrovic, Osborne, and Lavrenko 2011; Suh et al. 2010). The [CLS] embedding is passed through a linear layer and is concatenated with sparse features: the log-scaled number of followers and existence of a URL. This concatenated embedding is passed through two additional layers to be transformed into a single scalar value, where we apply a sigmoid function to convert into a score between 0 and 1, with 1 indicating a potential retweet.

Relationship-aware Model extends our baseline model using a number of representations for relationship information. We first use a direct encoding of the relationship category into a 256-dimensional vector which is trained along with other parameters. The textual information of the declared phrase associated with the relationship type (e.g., "my *best friend*") is also encoded using a character-level CNN with 1-d convolution (Kim 2014) then max-pooled, resulting in another representation vector of size 768. This number is obtained by using 256 convolutional filters each for kernels with sizes of 3,4, and 5. Finally, we add the relationship category as one-hot features in addition to the other sparse features, and concatenate them with the dense embeddings.

We set d to 768 and the learning rate to 1e-6. All models are trained with batch sizes of 16 and for a maximum of 10 epochs. We select the model with the highest validation F1 score, which is computed for every 5000 steps.

6.3 Results

Adding relationship information improves performance for retweet prediction, as shown in Table 5, which is known to be a difficult task (Martin et al. 2016). Incorporating relationship types provides a 1% performance increase in F1 score due to an increase in recall by 2%. Here, we show separate perfor-
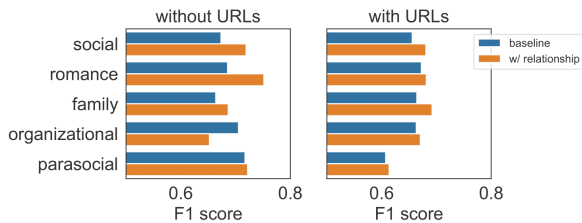

Figure 5: A comparison of baseline vs. relationship-infused models for retweet prediction on tweets with and without URLs. In both settings, the addition of a relationship type improves the predictive performance.

mances for test data tweets with and without URLs; tweets with URLs are likely to have different retweet dynamics on the basis of the content in the URL (e.g., retweeting a linked news story versus a personal message) and the different social uses of retweets (boyd, Golder, and Lotan 2010).

Analyzing performance improvement by relationship type reveals a more complex picture of improvement, shown in Figure 5. For tweets containing URLs, the addition of relationship information consistently improves performance, while we observe the largest performance gains for tweets without URLs. In particular, in tweets without URLs, the model sees increases of 2.2%, 3.1% and 2.3% for social, romance, and family categories respectively, signalling that the model can use this social information to decide where a person in that relationship is likely to retweet based on the content. However, the model performs worse for predicting retweets from organizational relations (2.8% decrease) which lowers the overall performance reported in Table 5. In the case of tweets containing URLs, F1 scores increase across all categories, with social (3.8%) and family (4.2%) categories benefiting from the largest gains.

We observe that the larger gains seen in social, romance, and family relationships are due to increased recall (see Table 4 in Supplemental Material). This increase suggests that individuals embedded in social, romance, and family relationships retweet content that is less likely to be retweeted normally (e.g., mundane personal events) because of the nature of the relationship, which the relationship-aware model is able to use to correctly identify the content will be retweeted. This result is further evidence of the interaction between communication patterns and relationship types.

7 Ethics and Limitations

Ethical Considerations This research was performed on only public data, in accordance with the Twitter Terms of Service. However, users do not necessarily expect their data to be collected for research purposes nor to be disclosed (Fiesler and Proferes 2018). Furthermore, the potentially-sensitive nature of interpersonal relationships being revealed necessitates that additional privacy steps must be taken and that benefits must outweigh harms. To mitigate risk to individuals, we report only aggregate information and focus on broad effects, avoiding any focus on marginalized groups or sensitive relationships (Townsend and Wallace 2016). Further,

data and models will only be shared upon confirmation of ethical principles of use. Counterbalancing these risks, this study offers substantial benefit to our understanding of social processes and how relationships influence what we share and what we hear about. As demonstrated in Section 6, understanding these relationships can improve the algorithms that individuals come into contact with regularly, such as social and content recommendation systems.

Limitations Data used in this study relies on self-reported, public declarations, which may not occur across all dyads in Twitter. The willingness to declare such relationships is likely indicative of a stronger tie between users. Therefore, our work may not reflect the behaviors of people in relationships less likely to be declared. Such non-declarations can be due to a variety of reasons such as increased desire for privacy, weaker association of that relationship type, or even the potential social stigma around declaring the relationship. Further, our data and model depend on observing communication between two users; as not all users with a meaningful social relationship also communicate on Twitter (e.g., spouses on Twitter who talk offline), our model is unable to identify such relationship. While the results of our work are largely in line with prior expectations from sociology and psychology, future work is needed to understand what biases, if any, stem from using only self-reported relationships and how these relationships fit within the broader space of relationships that exist between users on Twitter. Finally, our validation efforts are at the population level and future work could perform additional validation on specific relationships through crowd-sourced labeling of both the self-declared and the inferred relationships obtained through our model.

Our analysis is based on a 10% sample of Twitter data and while consisting of billions of tweets, this sampling by nature omits tweets between users that would affect the inference of network and communication statistics. Specifically, while the measured using the best-available data and follow past scholarship that used similar datasets to infer network properties (e.g., Bliss et al. 2014; Pierri, Piccardi, and Ceri 2020), our results likely underestimate the presence of edges and rates of communication frequency due to sampling.

The current work makes two simplifying assumptions about a dyad's relationships: (i) the relationship is of only one type and (ii) does not change over time. In practice, relationships evolve over time and categories such as social and organizational can indeed overlap. Our simplifying assumptions allow us to perform these initial studies at large scale. However, future work could relax these constraints with sufficient longitudinal data or with additional self-reported data.

8   Related Work

Due to the difficulty of collecting ground truth data, only a handful of studies have examined real social relationship of different categories on social media. Min et al. (2013) conduct a survey on 40 participants to obtain their SMS data and categorize their contacts into *family*, *work*, and *social*. The authors show that it is possible to infer these relationship types using features such as the geographical similarity between two users and their contact patterns. More recently, Welch et al. (2019) collected private data for the ego network of one user and 104 alters where there exact relationship was known; the combination of text and behavioral data (such as message duration) were used to predict relationship properties. Other studies have predicted the existence of a specific relationship from interpersonal interactions, the most common being a romantic relationship. Backstrom and Kleinberg (2014) show that dispersion—a metric which the authors introduce for measuring how well-connected mutual neighbors are—is a better predictor of romantic relationships than embeddedness within the Facebook network. Similarly, Tay, Tuan, and Hui (2018) introduce a model that compares the similarity of messages shared between two people on Twitter to predict whether they are in a romantic relationship. Our study goes substantially beyond these studies by simultaneous testing a comprehensive set of relationship types and examining orders of magnitude more data for each relationship type.

Rather than study categories of relationships, some studies instead measure attributes of the relationship themselves, e.g., their relative status. Adali, Sisenda, and Magdon-Ismail (2012) use social and behavioral features derived from Twitter activities to distinguish between relationship groups defined by their word usage and a clustering algorithm. Gilbert and Karahalios (2009) and Gilbert (2012) use interaction-based features provided by the Facebook platform to predict the tie strength of user dyads in Facebook and Twitter. Rashid and Blanco (2017) and Rashid and Blanco (2018) use a conversation dataset from the TV series *Friends* to label the relationships between the main characters with properties such as "equal-hierarchical" or "pleasure-task oriented", then show that these properties can be predicted using text features. Choi et al. (2020) predict the prevalence of specific social dimensions such as trust or power differences using deep learning classifiers on Twitter interactions between two users. These studies offer a complementary view of relationships and our introduction of a new large-scale dataset with relationship categories opens up new future work for testing how these attributes align or differ across categories.

Social networking platforms serve multiple roles as both social and informational networks (Arnaboldi et al. 2013), and as such individuals may form ties with others for different purposes. Several works have examined how different properties of the dyads reflect their information sharing behavior. In particular, studies have focused on the characteristics of interactions between close users, showing that such dyads (i) have a greater tendency to share less common hashtags which may indicate community belonging (Romero, Tan, and Ugander 2013), (ii) are more likely to share content than with strangers (Quercia, Capra, and Crowcroft 2012; Bakshy, Karrer, and Adamic 2009), and (iii) frequently engage through actions such as mentioning each other or sharing posts (Jones et al. 2013). Other work focuses on user interactions with influential users such as celebrities or politicians who possess a large follower base. These users gain widespread influence by specializing on narrow topics (Cha et al. 2010) or posting messages with strong sentiments (Dang-Xuan et al. 2013). While our study also aims to identify communication and interaction differences between different types of user rela-

tionships, we go beyond the widely studied themes, such as close ties or influential users, to study how specific types of relationships interact with information sharing, showing differences in which content is communicated across specific types (e.g., Figure 1) and that knowing the relationship type aid in predicting which content is retweeted.

9 Conclusion

Not all ties are equal: friends, family, and lovers all have different social, linguistic, and temporal behaviors—yet, social network studies have typically limited themselves to networks with edges that encode only the existence of a relationship, but ignore the *type* of that relationship. Using a dataset consisting of the interactions between 9.6M dyads on Twitter with known relationship types, we introduce a new approach that explicitly models interpersonal relationship types in social networks. We make the following three contributions towards understanding relationships in networks. First, we show that the linguistic, topical, network, and diurnal properties in online communication between different relationship types match predictions from theory and observational studies. Second, we demonstrate that relationship types can be accurately predicted using text and network features combined with state-of-the-art deep learning models. Third, we show that knowing relationship types improves performance when predicting retweets—demonstrating the benefits of predicting relationships at scale. The addition of relationship type significantly improves the recall of retweets for social, family, and romance relationships, which are considered more personal.

Our proposed approach, combined with the consistency of our results with existing literature on social relationships, further demonstrates the value in studying social media networks to further understand the differences in communicative behaviour across interpersonal relationships. Furthermore, as evident from the performance of our relationship classifier and the improvement on retweet prediction, our work enables new types of analyses that benefit from large-scale relationship-aware networks such as modeling network evolution, information diffusion dynamics, and community structure. Overall, our work provides a stepping stone towards incorporating relationship types in several research questions in social and network sciences.

Acknowledgments

References

Adali, S.; Sisenda, F.; and Magdon-Ismail, M. 2012. Actions Speak as Loud as Words: Predicting Relationships from Social Behavior Data. In *WWW*.

Adamic, L. A.; and Adar, E. 2003. Friends and neighbors on the Web. *Social Networks* 25(3): 211 – 230.

Altman, I.; and Taylor, D. A. 1973. *Social penetration: The development of interpersonal relationships.* Holt, Rinehart & Winston.

Arnaboldi, V.; Conti, M.; Passarella, A.; and Dunbar, R. 2013. Dynamics of Personal Social Relationships in Online Social Networks: A Study on Twitter. In *COSN*.

Backstrom, L.; and Kleinberg, J. 2014. Romantic Partnerships and the Dispersion of Social Ties: A Network Analysis of Relationship Status on Facebook. In *CSCW*.

Bakshy, E.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Everyone's an Influencer: Quantifying Influence on Twitter. In *WSDM*.

Bakshy, E.; Karrer, B.; and Adamic, L. A. 2009. Social influence and the diffusion of user-created content. In *EC*.

Bell, A. 1984. Language style as audience design. *Language in society* 13(2): 145–204.

Beller, C.; Knowles, R.; Harman, C.; Bergsma, S.; Mitchell, M.; and Van Durme, B. 2014. I'm a Belieber: Social Roles via Self-identification and Conceptual Attributes. In *ACL*.

Bergsma, S.; and Van Durme, B. 2013. Using Conceptual Class Attributes to Characterize Social Media Users. In *ACL*.

Berscheid, E.; Snyder, M.; and Omoto, A. M. 1989. The Relationship Closeness Inventory: Assessing the Closeness of Interpersonal Relationships. *Journal of Personality and Social Psychology* 57(5): 792–807.

Bliss, C. A.; Frank, M. R.; Danforth, C. M.; and Dodds, P. S. 2014. An evolutionary algorithm approach to link prediction in dynamic social networks. *Journal of Computational Science* 5(5): 750 – 764.

boyd, d.; Golder, S.; and Lotan, G. 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *HICSS*.

Bridge, K.; and Baxter, L. A. 1992. Blended relationships: Friends as work associates. *Western Journal of Communication* 56(3): 200–225.

Burke, M.; Adamic, L.; and Marciniak, K. 2013. Families on Facebook. In *ICWSM*.

Burke, M.; and Kraut, R. E. 2014. Growing Closer on Facebook: Changes in Tie Strength through Social Network Site Use. In *CHI*.

Burton-Chellew, M. N.; and Dunbar, R. I. 2015. Romance and reproduction are socially costly. *Evolutionary Behavioral Sciences* 9(4): 229–241.

Cha, M.; Haddadi, H.; Benevenuto, F.; and Gummadi, K. 2010. Measuring user influence in twitter: The million follower fallacy. In *ICWSM*.

Chen, T.; and Guestrin, C. 2016. XGBoost: A Scalable Tree Boosting System. In *KDD*.

Cheng, J.; Romero, D. M.; Meeder, B.; and Kleinberg, J. 2011. Predicting reciprocity in social networks. In *SocialCom*.

Choi, M.; Aiello, L. M.; Varga, K. Z.; and Quercia, D. 2020. Ten Social Dimensions of Conversations and Relationships. In *WWW*.

Choudhury, M. D.; and De, S. 2014. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. In *ICWSM*.

Colleoni, E.; Rozza, A.; and Arvidsson, A. 2014. Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication* 64(2): 317–332.

Dang-Xuan, L.; Stieglitz, S.; Wladarsch, J.; and Neuberger, C. 2013. An investigation of influentials and the role of sentiment in political communication on Twitter during election periods. *Information, communication & society* 16(5): 795–825.

David-Barrett, T.; Kertesz, J.; Rotkirch, A.; Ghosh, A.; Bhattacharya, K.; Monsivais, D.; and Kaski, K. 2016. Communication with Family and Friends across the Life Course. *PLOS ONE* 11(11): 1–15.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805* .

Ellison, N. B.; Gray, R.; Vitak, J.; Lampe, C.; and Fiore, A. T. 2013. Calling All Facebook Friends: Exploring Requests for Help on Facebook. In *ICWSM*.

Ellison, N. B.; Steinfield, C.; and Lampe, C. 2007. The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication* 12(4): 1143–1168.

Fägersten, K. B. 2012. *Who's swearing now?: the social aspects of conversational swearing*. Cambridge Scholars Publishing.

Farnham, S. D.; and Churchill, E. F. 2011. Faceted Identity, Faceted Lives: Social and Technical Issues with Being Yourself Online. In *CSCW*.

Fast, E.; Chen, B.; and Bernstein, M. S. 2016. Empath: Understanding Topic Signals in Large-Scale Text. In *CHI*.

Feld, S. L. 1981. The Focused Organization of Social Ties. *American Journal of Sociology* 86(5): 1015–1035.

Feld, S. L. 1982. Social Structural Determinants of Similarity among Associates. *American Sociological Review* 47(6): 797–801.

Feldman, G.; Lian, H.; Kosinski, M.; and Stillwell, D. 2017. Frankly, We Do Give a Damn: The Relationship Between Profanity and Honesty. *Social Psychological and Personality Science* 8(7): 816–826.

Fiesler, C.; and Proferes, N. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4(1).

Foster, J. G.; Foster, D. V.; Grassberger, P.; and Paczuski, M. 2010. Edge direction and the structure of networks. *Proceedings of the National Academy of Sciences* 107(24): 10815–10820.

Garimella, K.; Cohen, J.; and Weber, I. 2017. Understanding Para Social Breakups on Twitter. In *WebSci*.

Gilbert, E. 2012. Predicting Tie Strength in a New Medium. In *CSCW*.

Gilbert, E.; and Karahalios, K. 2009. Predicting Tie Strength with Social Media. In *CHI*.

Golder, S. A.; and Macy, M. W. 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science* 333(6051): 1878–1881.

Gorrese, A.; and Ruggieri, R. 2012. Peer Attachment: A Meta-analytic Review of Gender and Age Differences and Associations with Parent Attachment. *Journal of Youth and Adolescence* 41(5): 650–672.

Granovetter, M. 1983. The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory* 1: 201–233.

Granovetter, M. S. 1973. The strength of weak ties. *American Journal of Sociology* 78(6): 1360–1380.

Hartup, W. W.; Shantz, C.; Eisenberg, N.; Emde, R.; Hoffman, L.; Maccoby, E. E.; Monks, F. J.; Parke, R.; Rutter, M.; and Zahn-Waxler, C. 1999. *The development of romantic relationships in adolescence*. Cambridge University Press.

Hartup, W. W.; and Stevens, N. 1999. Friendships and Adaptation Across the Life Span. *Current Directions in Psychological Science* 8(3): 76–79.

Hays, R. B. 1984. The development and maintenance of friendship. *Journal of Social and Personal Relationships* 1(1): 75–98.

Hong, L.; Dan, O.; and Davison, B. D. 2011. Predicting Popular Messages in Twitter. In *WWW*.

Huang, B.; and Carley, K. 2019. A Hierarchical Location Prediction Neural Network for Twitter User Geolocation. In *EMNLP*.

Jenders, M.; Kasneci, G.; and Naumann, F. 2013. Analyzing and predicting viral tweets. In *WWW*.

Jones, J. J.; Settle, J. E.; Bond, R. M.; Fariss, C. J.; Marlow, C.; and Fowler, J. H. 2013. Inferring Tie Strength from Online Directed Behavior. *PLOS ONE* 8(1): 1–6.

Kehrberg, A. K. 2015. 'I love you, please notice me': the hierarchical rhetoric of Twitter fandom. *Celebrity Studies* 6(1): 85–99.

Kim, Y. 2014. Convolutional Neural Networks for Sentence Classification. In *EMNLP*.

Kirilenko, A. P.; and Stepchenkova, S. 2016. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLOS ONE* 11(3): 1–14.

Klein, K. J.; and D'Aunno, T. A. 1986. Psychological sense of community in the workplace. *Journal of Community Psychology* 14(4): 365–377.

Knapp, M. L.; Ellis, D. G.; and Williams, B. A. 1980. Perceptions of communication behavior associated with relationship terms. *Communication Monographs* 47(4): 262–278.

Kossinets, G.; and Watts, D. J. 2006. Empirical Analysis of an Evolving Social Network. *Science* 311(5757): 88–90.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? In *WWW*.

Lee, S. J. 2009. Online Communication and Adolescent Social Ties: Who benefits more from Internet use?*. *Journal of Computer-Mediated Communication* 14(3): 509–531.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010a. Predicting Positive and Negative Links in Online Social Networks. In *WWW*.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010b. Signed Networks in Social Media. In *CHI*.

Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; and Stoyanov, V. 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *CoRR* abs/1907.11692.

Martin, T.; Hofman, J. M.; Sharma, A.; Anderson, A.; and Watts, D. J. 2016. Exploring Limits to Prediction in Complex Social Systems. In *WWW*.

Marwell, G.; and Hage, J. 1970. The Organization of Role-Relationships: A Systematic Description. *American Sociological Review* 35(5): 884–900.

McCallum, A. K. 2002. MALLET: A Machine Learning for Language Toolkit. http://mallet.cs.umass.edu. Accessed: 2021-04-10.

Min, J.-K.; Wiese, J.; Hong, J. I.; and Zimmerman, J. 2013. Mining Smartphone Data to Classify Life-Facets of Social Relationships. In *CSCW*.

Myers, S. A.; and Johnson, A. D. 2004. Perceived solidarity, self-disclosure, and trust in organizational peer relationships. *Communication Research Reports* 21(1): 75–83.

Ozenc, F. K.; and Farnham, S. D. 2011. Life "Modes" in Social Media. In *CHI*.

Pennebaker, J. W.; Boyd, R. L.; Jordan, K.; and Blackburn, K. 2015. The development and psychometric properties of LIWC2015. https://repositories.lib.utexas.edu/handle/2152/31333. Accessed: 2021-04-10.

Petrovic, S.; Osborne, M.; and Lavrenko, V. 2011. RT to Win! Predicting Message Propagation in Twitter. In *ICWSM*.

Pierri, F.; Piccardi, C.; and Ceri, S. 2020. Topology comparison of Twitter diffusion networks effectively reveals misleading information. *Scientific reports* 10(1): 1–9.

Quercia, D.; Capra, L.; and Crowcroft, J. 2012. The Social World of Twitter: Topics, Geography, and Emotions. In *ICWSM*.

Ranganath, S.; Tang, J.; Hu, X.; Sundaram, H.; and Liu, H. 2015. Leveraging Social Foci for Information Seeking in Social Media. In *AAAI*.

Rashid, F.; and Blanco, E. 2017. Dimensions of interpersonal relationships: Corpus and experiments. In *EMNLP*.

Rashid, F.; and Blanco, E. 2018. Characterizing Interactions and Relationships between People. In *EMNLP*.

Richey, M. H.; and Richey, H. W. 1980. The significance of best-friend relationships in adolescence. *Psychology in the Schools* 17(4): 536–540.

Rivas, J. 2009. Friendship selection. *International Journal of Game Theory* 38(4): 521.

Roberts, S. G. B.; and Dunbar, R. I. M. 2011. Communication in social networks: Effects of kinship, network size, and emotional closeness. *Personal Relationships* 18(3).

Romero, D. M.; Tan, C.; and Ugander, J. 2013. On the Interplay between Social and Topical Structure. In *ICWSM*.

Shulman, N. 1975. Life-Cycle Variations in Patterns of Close Relationships. *Journal of Marriage and Family* 37(4).

Sias, P. M.; Gallagher, E. B.; Kopaneva, I.; and Pedersen, H. 2012. Maintaining Workplace Friendships: Perceived Politeness and Predictors of Maintenance Tactic Choice. *Communication Research* 39(2): 239–268.

Sluss, D. M.; and Ashforth, B. E. 2007. Relational Identity and Identification: Defining Ourselves through Work Relationships. *The Academy of Management Review* 32(1).

Stafford, L.; and Canary, D. J. 1991. Maintenance Strategies and Romantic Relationship Type, Gender and Relational Characteristics. *Journal of Social and Personal Relationships* 8(2): 217–242.

Stever, G.; and Lawson, K. 2013. Twitter as a Way for Celebrities to Communicate with Fans: Implications for the Study of Parasocial Interaction. *North American Journal of Psychology* 15(2): 339–354.

Suh, B.; Hong, L.; Pirolli, P.; and Chi, E. H. 2010. Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. In *SocialCom*.

Tang, J.; Wang, X.; and Liu, H. 2012. Integrating Social Media Data for Community Detection. In Atzmueller, M.; Chin, A.; Helic, D.; and Hotho, A., eds., *MSM*.

Tay, Y.; Tuan, L.; and Hui, S. 2018. CoupleNet: Paying Attention to Couples with Coupled Attention for Relationship Recommendation. In *ICWSM*.

Townsend, L.; and Wallace, C. 2016. Social media research: A guide to ethics. The University of Aberdeen.

Tukachinsky, R. 2010. Para-romantic love and para-friendships: Development and assessment of a multiple-parasocial relationships scale. *American Journal of Media Psychology* 3(1/2): 73–94.

Tuulos, V. H.; and Tirri, H. 2004. Combining Topic Models and Social Networks for Chat Data Mining. In *WI*.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. In *NeurIPS*.

Verbrugge, L. M. 1983. A Research Note on Adult Friendship Contact: A Dyadic Perspective. *Social Forces* 62(1): 78–83.

Wang, Y.-C.; Burke, M.; and Kraut, R. 2016. Modeling Self-Disclosure in Social Networking Sites. In *CSCW*.

Welch, C.; Pérez-Rosas, V.; Kummerfeld, J. K.; and Mihalcea, R. 2019. Look Who's Talking: Inferring Speaker Attributes from Personal Longitudinal Dialog. In *CICLing*.

Yuan, N. J.; Zhong, Y.; Zhang, F.; Xie, X.; Lin, C.-Y.; and Rui, Y. 2016. Who Will Reply to/Retweet This Tweet?: The Dynamics of Intimacy from Online Social Interactions. In *WSDM*.