Modeling Social Support Scores Using Phone Use Patterns

Isha Ghosh
Rutgers University,
New Brunswick, NJ.
Isha.ghosh@rutgers.edu

Vivek K. Singh
Rutgers University and Massachusetts Institute of Technology,
New Brunswick, NJ.
vivek.k.singh@rutgers.edu

ABSTRACT

Social support is a fundamental concept of interactional behavior and has been consistently linked with higher levels of emotional and physical well-being. Social support has been shown to be an effective coping mechanism in managing stress and depression and to provide positive experiences that directly enhance overall well-being and quality of life. The measurement and quantification of social support has generated a great deal of interest in the behavioral science community over the past 10 years. The dominant approach for quantifying individual social support remains self-reported surveys and generator-methods, which are costly, attention-consuming, and fraught with biases. Given the important role played by mobile phones in mediating human social lives, this study explores the use of phone metadata (call and SMS logs) to automatically infer an individual's social support. Based on Sherbourne and Stewart's Modified Social Support survey (MSSS) as ground truth and ten-week phone data collection for 55 participants, we report that multiple call and SMS based social features are intrinsically associated with social support and perform better than demography based models.

KEYWORDS

Social Informatics, Phone Mediated Communication, Data Science

INTRODUCTION

Social support is defined as the resources or aids exchanged between individuals through interpersonal ties (Cohen & Willis, 1985) and the causes and effects of social support have stimulated research for almost 30 years (Dunkel-Schetter & Brooks, 2009; Wortman & Dunkel-Schetter, 1987). Access to social support has been shown to trigger many positive outcomes such as wellbeing and life satisfaction (Turner & Brown, 2010) as well as decrease the negative effects of physical and psychological illnesses more than any other buffering factor (Cohen & Wills, 1985). Social support is often seen as an interactional behavior through which individuals express, and receive emotional concern, instrumental aid, or information (Dunkel-Schetter & Brooks, 2009). Different aspects of an individuals' social network (online or offline), for example the size and diversity of the network, characteristics of their ties, (e.g. frequency of contact, multiplexity, and tie strength), and the attributes of the individual network members (e.g. age and gender) have been shown to have a relationship with their social support (Wellman & Frank, 2001). Social support consists of a variety of helping behaviors performed by one person for the benefit of another. Examples of such behaviors include giving advice, empathizing, assisting with practical tasks, and expressing encouragement (Barrera, Sandler, & Ramsay, 1981). Both receiving support, and perceiving that support is available, are associated with positive outcomes. Numerous studies indicate that people with spouses, friends, and family members who provide psychological and material resources are in better health than those with fewer supportive social contacts. Having access to a network that provides social support has been linked to reduction of stress, improvement of mental-health and depression, and increasing the overall well-being experienced by individuals (Haslam, O'brien, Jetten, Vormedal, & Penna, 2005).

Previous research on how social support is transferred and received in social networks often builds on the theoretical concept of social capital (Bourdieu, 1980; Putnam, 1995). Williams (2006) adapted the concept of social capital to describe social support within online contexts and developed an operationalization of online social capital. Community ties with friends and relatives often provide different forms of social support for individuals. These relationships have also been shown to translate into social capital that people use to deal with daily life, seize opportunities, and reduce uncertainties (Kadushin, 1981). Social support needs to be differentiated from social capital, though. Social capital refers to the resources from one's social network (Lin, 2002) and originated in sociological research, whereas social support is a psychological construct. Adler and Kwon (2002) suggest to clearly distinguish between the source and the effects of social capital. The source is "the structure and content of the actor's social relations", whereas the effects are "the information, influence, and solidarity it makes available to the actor" (p. 23). Accordingly, social support is a possible effect of social capital.

Recent research investigating access to social support from online sources has highlighted how informational and emotional support is transacted in online settings (Ellison, Vitak, Gray, & Lampe, 2014; Chang, 2009). Informational support shows strong connections to the concept of bridging social capital as weak ties offer access to a wide array of connections, which may be capitalized in terms of novel information (Ellison et al., 2014). Emotional support is strongly linked to bonding social capital, as strong ties are an important source of emotional support. While the context-dependent aspects of social support have been previously studied, modeling the overall social support accessible to a person from various aspects of their network remains an open area of research.

Studies on online social networking have examined the exchange of social support through online social networking and its outcomes. Acquiring social support from others in a social network is found to be one of the most important reasons for online social networking (Oh & Syn, 2015; Oh, Ozkaya, & LaRose, 2013). People use online spaces to discuss problems or obtain information that is helpful when coping with particular stressors. Social media sites may fulfill a need for social belongingness, distract people from various stressors, or offer micro-boosts to self-esteem by being "friended," "liked," or "followed" by others (Cole, Nick, Zelkowitz, Roeder, & Spinelli, 2017). Previous studies from several different disciplines including communication, psychology and sociology suggest that the Internet offers opportunities to exchange information and consolation with online associates (Salehan & Negahban, 2013; Walther & Boyd, 2002) and fosters meaningful relationships (Parks & Floyd, 1996). Digital communications have also shown to help in mobilizing social support as well as maintaining and strengthening existing relationships with geographically near and distant contacts (Quan-Haase, Mo, & Wellman, 2017). While this research addresses the relationship between internet social network communities and the creation and maintenance of social support, the study of various facets of mobile phone usage as a facilitator of supportive social networks remains an open question.

Traditional approaches for quantifying individual social support have focused on aspects that could be simply observed (e.g., gender, race, age) or elicited in a small period of time in lab settings (e.g., via surveys and generator methods) (Gottlieb & Bergen, 2010). Each of these approaches relies on active human reporting and must contend with numerous hurdles such as subjectivity in reports and observations, social and cognitive biases, and narrow observation chances, while dealing with pressures such as budget, time, and the effort required. *This implies that there are no current approaches that infer an individual's social support levels without the need for active human effort.*

In today's world, mobile phones have become one of the most favored devices to interact and communicate with friends, family, as well as engage with digital information in its various forms. As interactions get increasingly mediated via mobile phones, individuals communicating with each other via these devices, form their own social network (Beale, 2005). Data generated from smartphones can provide insights into how often and which members of the network people choose to communicate with, changes in conversational patterns, as well as length and breadth of communication with different members of the network. Recently, mobile phones along with sensor-based data have been used by multiple researchers to construct rich and individualized models of human behavior in social, spatial, and temporal settings, and link them to individual personality traits such as privacy attitudes and trust propensity (Singh & Ghosh, 2017; Blumenstock, Cadamuro, & On, 2015). Emerging research has explored the potential of using information passively (i.e. without active human effort) gathered from these phone-mediated interactions to make predictions about an individual's economic status, personality traits, and well-being (de Montjoye, Quoidbach, Robic, & Pentland, 2013). Given such recent results and the theoretical literature connecting social support with social capital, mental health and well-being, (Ruvalcaba-Romero, Fernández-Berrocal, Salazar-Estrada, & Gallegos-Guajardo, 2017), this study explores the creation of an automated phone data-based approach for modeling individual social support. Such a phone-based method, if successful, would offer a low-cost, faster, scalable, and automatic method for generating insights into social support for millions of users with applications in information science research, health and wellness, as well as communications and sociology. Specifically, this exploratory study examines the following research questions:

RQ1. Which (if any) long-term phone-use patterns have associations with an individual's social support levels?
RQ2. Can a machine learning algorithm be used to automatically infer individual social support based on phone metadata?

In this work, we analyze data from a ten-week field study to systematically study the interconnections between phone-based behavioral measures (e.g. number of phone calls made) and "ground truth" modified social support survey scores based on Sherbourne and Stewart's (1991) survey for 55 individuals.

STUDY

We study the interconnections between social support and phone-based features on the data gathered as part of <Anonymized> Study undertaken at a major university in the Northeastern part of the United States of America. The participants for the study were recruited using flyers, email announcements, and social media posts in the area surrounding a major North American university. The participants needed to: (1) be between 18-75 years of age; (2) comfortable with written and oral English; (3) use an Android smartphone; (4) carry their phone on them most of the time; and (5) be willing and able to travel to the study site for three in-person sessions. This can hence be considered a convenience sample, which was considered acceptable, given the exploratory nature of this work.

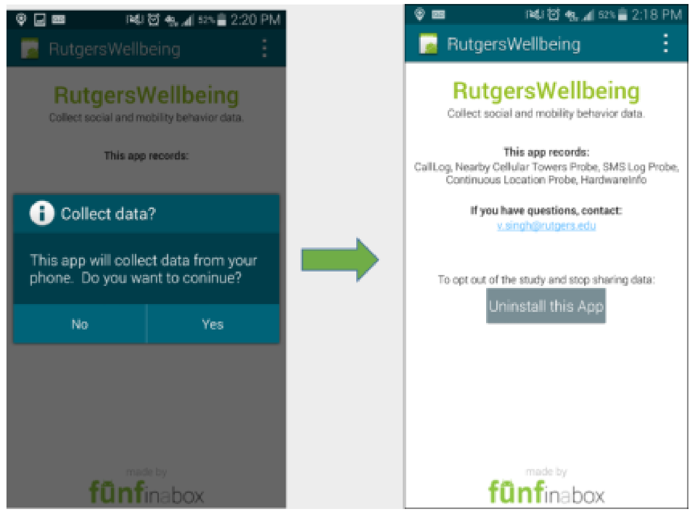

Figure 1. Screenshot of the App (anonymized for review)

Participants in this study were asked to attend three in-person sessions for surveys and to install a mobile application (Figure 1) onto their smartphone. The app recorded anonymized call and SMS metadata (calls/SMS initiated or received - number, times and anonymized id but no actual audio or text). The app also recorded location metadata, which is not relevant to the current discussion. The study included 59 participants. However, some of the participants did not complete all the surveys, and some did not enter their unique identifying code consistently across different surveys, resulting in a set of 55 participants for whom we have the mobile-based data as well as the scores for the two surveys of interest (more details on surveys presented later). The survey order was randomized for different participants. Participation in this study was voluntary and incentivized monetarily. The participants were compensated up to a sum of $100 on successful completion of the study over the ten-week period.

MEASURES

Demographic Descriptors

The sample consisted of 36 male and 19 female participants (university population is 50% male, 50% female). The most common age group for participants was 18-21 years, the most common education level was “some college,” the median annual family income was in the range $35,000-$49,000 and 96% of them were single. A visual representation of the demographic is presented in Figure 2.

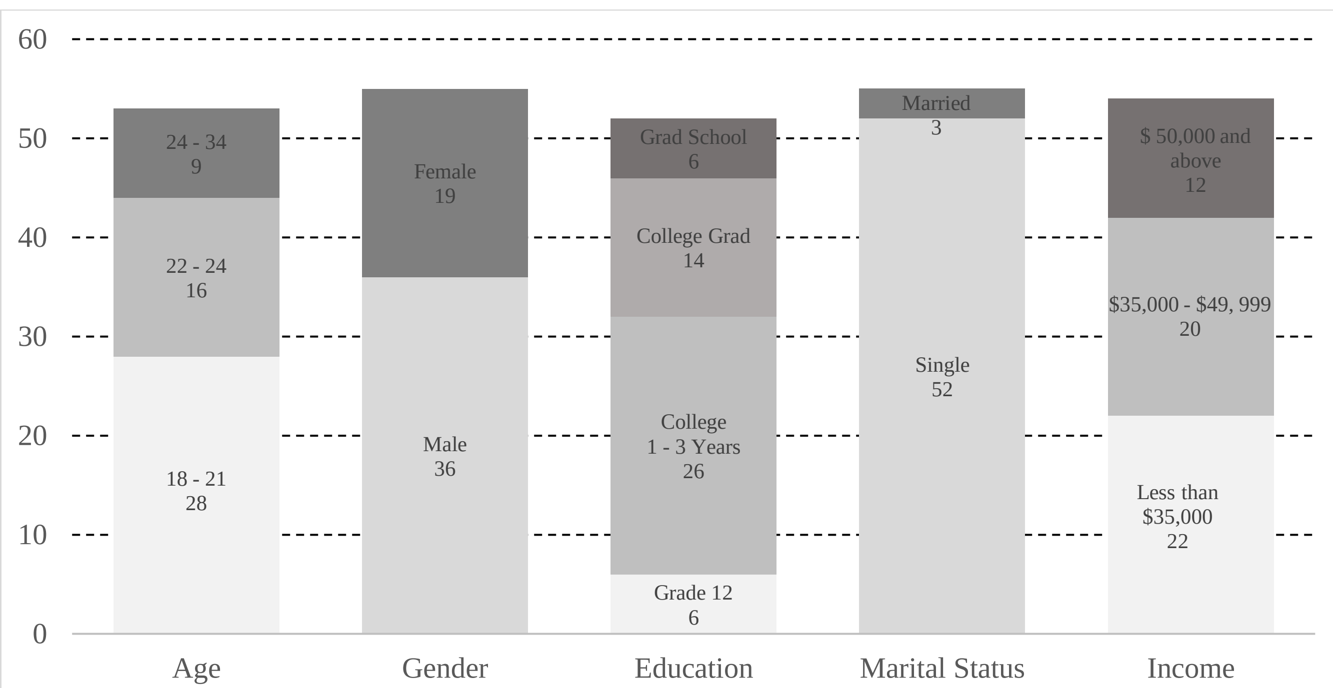

Figure 2. Descriptive statistics of the sample population

Data type	Number of participants	Data points
Calls	55	28,132 calls
SMS	55	187,720 messages
Surveys	55	2 surveys per participant.

Table 1. Summary of data considered in this study

A summary of the data collected over the ten-week period in Spring 2015 is shown in Table 1. The participants on average made 511 phone calls (median = 312; minimum = 43; maximum = 2,711) and exchanged 3,413 SMS messages (median = 2,423; minimum = 21; maximum = 17,625) over this period.

Social Support

The survey was based on Sherbourne and Stewart's Modified Social Support Survey (1991). This is a 19-item, self-administered social support survey measuring measure functional social support. It has been actively used by multiple studies (>4000 citations per Google Scholar) and is considered universally applicable (McDowell, 2006). The 19 items cover four domains of social support (emotional/informational support, tangible support, positive social interaction, and affection) as well as an overall social support score. Participants were asked to answer questions on a 5-point scale ranging from Strongly Disagree (1) to Strongly Agree (5).

Minimum	Maximum	Mean	Median
2	5	3.76	4

Table 2. Summary of Social Support Scores

A descriptive summary of the overall social support score is shown in Table 2.

While the different constructs of social support have previously been studied in relationship with online social networks, the quantification of the overall social support experienced by an individual remains to be studied. The unique affordances of a mobile phone (e.g.: access to synchronous and asynchronous modes of communication via call and text) make it the device of choice to connect with ties of different strengths in different modes (synchronous, asynchronous), across differing times and locations and provides different forms of social support.

Characterizing Phone-based Social Support

At a conceptual level, social support is closely connected with an individual's relative position in the network (Trepte, Dienlin, & Reincke, 2015). On a more granular level, the frequency of communication and access to strong and weak ties in a network has also been connected to social support (Oh et al., 2013). Such features have been operationalized over online social networks (Cole et al., 2017) and recently over phone networks in different contexts (Quan-Haase et al., 2015). At an empirical level, multiple recent efforts have quantified different aspects of individual behavior based on phone metadata with a goal to identify personal traits and well-being (de Montjoye et al., 2013; Singh & Ghosh, 2017).

Based on a survey of existing literature we characterize phone use features into four categories to understand different aspects of an individual's social behavior. In defining these features (total 17), we integrate some of the more obvious usage patterns (total number of calls/sms messages) with more nuanced/exploratory aspects of phone usage while keeping the number of features manageable.

Type	Phone Features	Empirical Support	Conceptual Framework
Social Activity	Call_Count Sms_Count Total_call_duration InOut_ratioCalls InOut_ratioSMS Call_Response_rate Missed_call_percentage	Oh et al., 2013 Utz & Breuer, 2016 Kovanen, Saramaki, & Kaski, 2010	Relational Investment Social Connectedness Donath, 2008; Ellison et al., 2014
Tie Strength	Call_Strong_ties; Sms_Strong_ties; Call_StrongWeak_ties_ratio; Sms_StrongWeak_ties_ratio	Wellman & Wortley, 1990 Luarn, Kuo, Chiu, & Chang, 2015 Utz & Breuer, 2016	Tie-Strength Granovetter (1973) Wellman, 1992
Relationship Maintenance	Call_Loyalty; Sms_Loyalty	Wellman, 1992	Relationship Maintenance Behavior Barrera, 1986

Temporal Rhythms	Day_night_Call_ratio; Day_night_sms_ratio; Weekday_weekend_call_ratio; Weekday_weekend_SMS_ratio	Jumin, Hyungeun, & Jung-hee, 2010 O'Keefe, & Sulanowski, 1995	Circadian Rhythms Abdullah et al., 2014; deMontjoye et al., 2013

Table 3. Summary of features considered in this study.

Social Activity

Theoretically, the amount of supportive interaction an individual has is associated with the size and frequency of communication within their network (Oh et al., 2013). In the context of a phone based network, we quantify social activity as the frequency and duration of interactions within a network. Social support is also defined as an "interpersonal transaction" (Utz & Breuer, 2016), this implies that reciprocity plays an important role in individuals' social activity. Here, ease of access is operationalized using features related to accepting or rejecting calls (Kovanen, Saramaki, & Kaski, 2010). In the context of phone use networks, social activity can be characterized by the following seven features: Call_Count; Sms_Count; Total_call_duration; InOut_ratioCalls; InOut_ratioSMS; Call_Response_rate; Missed_call_percentage.

Tie Strength

The strength of relationships in a network have been shown to have a significant impact on supportive ties (Wellman & Wortley, 1990). Prior literature also connects strong and weak ties with different aspects social support (Wellman, 1992). On social network sites, responding to a post via a comment or message has shown to be indicative of a strong tie and provide emotional support to users (Luarn, Kuo, Chiu, & Chang, 2015). The number of comments or "likes" received for a post has also been shown to be indicative of support (Utz & Breuer, 2016). Since frequency of interaction is an important predictor of tie strength, to quantify the relative role of different types of ties in one's social network we approximate "strong ties" as the top-third most frequent of their contacts. For both calls and SMS, we quantify the number of interactions that take place with these "strong ties" as well as the relative percentage of all interactions that take place with these contacts. We characterize the following four features as indicative of tie-strength: Call_Strong_ties; Sms_Strong_ties; Call_StrongWeak_ties_ratio; Sms_StrongWeak_ties_ratio

Relationship Maintenance

Humans are a social species, with many survival advantages accruing from our connections to others (Wellman, 1992). Because humans have a fundamental need to feel that they belong to a group, social interaction should improve well-being by fulfilling these needs. Satisfaction will be greatest when there is frequent communication within a subset of the network (Luarn et al., 2015). This is because the need to belong is not satisfied by social interaction alone, but also requires stable interpersonal relationships marked by positive concern and caring. The people an individual chooses to reach out to in times of need is also considered an important indicator of social support. We define "loyalty" as the percentage of communication (call/sms) occurring only with the top three contacts of an individual. A high loyalty score (close to 1) would indicate a tendency to prefer repeated interactions with a small number of favorite connections. We characterize the following two features as indicative of relationship maintenance within phone-based networks: Call_Loyalty; Sms_Loyalty

Temporal Rhythms

While mobile phones allow us to be always connected and available, prior research has documented a difference in the interactions made during phases of the day (e.g. morning, evening, night). Further, interactions during late night (past midnight) have been shown to have predictive power on mental health and well-being (Jumin, Hyungeun, & Jung-hee, 2010). Research has also shown a difference in the frequency and quality of phone calls made during weekdays and weekends (O'Keefe, & Sulanowski, 1995). Hence, to understand the rhythms in user activity i.e. consistency across different times of day and different days, we quantify temporal phone usage pattern as the ratio of calls/sms received during AM Phase (12 AM – 11:59 AM) to calls/sms received during the PM phase (12 PM – 11:59 PM) and the ratio of calls/sms received during weekdays to calls/sms received during weekend. We define the following four features as indicative of temporal rhythms: Day_night_Call_ratio; Day_night_sms_ratio; Weekday_weekend_call_ratio; Weekday_weekend_SMS_ratio.

RESULTS- INFERRING SOCIAL SUPPORT

Predictive Model

Predicting social support scores from phone use patterns level is a regression problem; that is, predicting an outcome variable (i.e., social support scores) from a set of input predictors (i.e., phone-based features). We utilize Lasso (Least Absolute Shrinkage and Selection Operator) regularized linear regression model as our predictive model (Tibshirani, 1996). This technique has previously been used to infer human values from social media text analysis (Chen, Hsieh, Mahmud, & Nicolas, 2014). Another study using data gathered from Android phones to understand how individuals handle apps stored on their phone also uses Lasso regression as a technique to shrink the number of features and create an efficient predictive model (Li, Ai, Liu, Tang, Huang, Feng, and Mei, 2016). Previous research shows that Lasso regression is most effective when we wish to shrink features or variables to enhance the efficiency of a predictive model. In the current study, we have a large number of behavioral features for a relatively small dataset. Given this previous research on the efficiency of Lasso regression, this technique was chosen to build the model. Lasso minimizes the sum of squared errors, with a bound on the sum of the absolute values of the coefficients. Thus, Lasso automatically selects more relevant features and discards redundant features. We have D = 17 input features and N = 55 training cases. Lasso has been specifically designed to help identify relevant features (i.e., predictors) in such settings (N not much greater than D), while avoiding overfitting.

We use Pearson's correlation coefficient (r) between the model's predicted value and the "ground truth" survey value to quantify the prediction performance. We also report the mean absolute error (MAE). A smaller MAE is preferred because it indicates that the predictions are closer to the ground truth. The MAE is in the same unit as the predicted variable and is hence relatively easy to interpret.

Prediction Results

We evaluate three different prediction models in this work. These correspond three different approaches (phone feature-based, demography-based, and phone + demography based). Demography based features have been connected with social support in numerous studies (e.g., Manalel & Antonucci, 2017) and hence we use this approach as a baseline. Specifically, the difference between the demography-based model and the phone-features based model helps us interpret the predictive power of phone-based features. Further, the combined phone + demography model quantifies the performance in scenarios where demography information is already available (e.g., to the phone company, automatically derived via phone features (Zhong, Tan, Mo, & Yang, 2013) or surveyed once to allow for repeated social support inferences over time.

We apply 15-fold cross validation to determine the parameters for Lasso and the weights for each feature in each evaluation. The results for the evaluation in terms of correlation coefficients between predicted and actual scores for the target variable are summarized in Table 4.

	Phone-features based model	Demography based model	Phone + Demography based model
Overall Social Support	0.65	0.21	0.68

Table 4. Correlation between the predicted values based on different models and the real values of overall support.

As can be seen in Table 4, the predicted social support strongly correlates with the ground truth with $r = 0.65$ ($p < 0.01$). As a baseline for comparison we also consider models that focus only on the available demographic features (age, gender, marital status, ethnicity, level of education, and family income level). While demographics explain some variance in the social capital levels, it can be clearly seen that phone-based features provide much higher predictive performance than that obtained by using only demography-based features. This underscores the value of phone-based features in automatically predicting social support scores for individuals. Lastly, the models that use both phone and demography based data have a correlation score of 0.68. This suggests that phone features are not merely replacements for demography-based features but rather, add complementary information.

	Phone-features based model	Demography based model	Phone + Demography based model
Overall Social Support	0.61	0.81	0.58

Table 5. Mean Absolute Error (MAE) for the predicted values based on different models and the real values of overall social support score.

We report the Mean Absolute Errors (MAE) between the actual and predicted values in Table 5. For instance, the MAE for predicted social support score using phone and demographic data was found to be 0.58, indicating that the predictions are within ±0.58 of the "ground truth" or survey based values of social support scores. Since the values for social support can theoretically range from 1 to 5, we consider an error of ±0.58 to be acceptable. We also note that the phone based models consistently yield lower errors than demography based models, and combining the phone and demography data yields the models with the lowest error rates.

Correlation Analysis

To understand the relative effect of different phone-based features on social support we undertook a *post-hoc* Pearson's correlation analysis between the social support scores obtained from the survey and phone-based features as shown in Table 6.

From Table 6, we see that the highest association was found with the total call duration ($r = 0.36$). This implies that the amount of time an individual spends on call interactions is more significant than the total number of calls for inferring a social support score. The calls and SMS made to strong ties (i.e. the top third most frequent contacts) also have a significant relationship with social support scores. Social support literature distinguishes the roles of strong and weak ties in providing social support. Strong ties are thought to provide empathic support, while weak ties maybe less willing to provide significant services, but do provide access to new opportunities and ideas (Granovetter, 1973). As Wellman (1992) notes, "Nurturing, caring, and tangible assistance are more likely to be expected and provided more readily in closer, multiplex, dense relationships, which also likely carry a presumption of reciprocity of supportive behavior". Research also argues that everyday support received from strong ties is what promotes well-being, confirming the recipients' sense of mattering to other people and sustaining a sense of self-worth (Utz & Breuer, 2016).

Phone Use features	Overall Social Support
Call_Count	0.082
Sms_Count	0.064
Total_Call_duration	0.364**
InOut_ratio	0.290*
Call_Response_rate	0.243
Missed_call_percentage	-0.258
Inout_ratio_SMS	0.218
Call_Loyalty	0.077
Sms_Loyalty	-0.198
Call_Strong_ties	0.343*
Sms_Strong_ties	0.276*
Call_SW_ties_ratio	0.246
Sms_SW_ties_ratio	-0.07
Day_night_Call_ratio	0.305*
Day_night_sms_ratio	-0.198
Weekday_weekend_call_ratio	-0.21
Weekday_weekend_SMS_ratio	0.204

Table 6. Bivariate correlation between phone-based features and overall social support score.

From the perspective of social support, it is just as important to initiate relationships as it is to be available to other members of the network. Similarly, we note that the ratio of incoming to outgoing calls is positively associated with social support. We also find a significant correlation between relative ratio of calls received after 12 am and social support scores (r = 0.30). A possible interpretation is that calls made or received late at night are likely to be from close family or friends, therefore, a higher number of such calls would imply that the individual had access to a larger supportive network. These results underscore the value of initiating calls, talking to multiple individuals in such calls, and access to contacts with whom one can interact even at less common hours (i.e. during AM phase.)

DISCUSSION

The first research question (RQ1) for this work aimed at understanding the associations (if any) between long-term phone-use patterns and an individual's social support. Based on the reasonably high explanatory power at a collective level (regression analysis) and multiple significant associations at an individual level (correlation analysis), we report that many of the phone-based features may indeed be intricately associated with individual social support.

The results of correlation analysis are largely in line with prior literature (Burke & Kraut, 2016). At the same time, they expand the understanding of these ties and the associations to how they manifest themselves over smartphones. The smartphone is a convenient, highly accessible, and capable device that is well suited to share and disclose information. It is a two-way device, creating and consuming information, is highly personal, and is almost always available. Based on Beale's (2005) work, this study treats people connected via the smartphone as part of a supportive social network and interactions performed within this network communicative of relationship creation and maintenance behavior.

Relationship maintenance is shown to be a significant factor in achieving a feeling of connectedness within a network and therefore, enhancing social support (Ellison et al., 2014). In addition to directly maintaining relationships though substantive communication, the symbolic value of communication also can maintain relationships independent of the content exchanged. For example, according to the theory of relational investment (Donath, 2008; Ellison et. al, 2014), the frequency and length of messages serve as signals of relationship value in social network sites. Donath (2008) states that "cost in time is a signal of the resources one is willing to commit to this relationship" (p. 238). According to this theory, more effortful communication should have a greater impact on social support, we find similar results for smartphone interactions. Research also points to embeddedness in a network as an important factor in gaining social support (Barrera, 1986). In the context of social networking sites, belonging within a network is demonstrated via explicit behaviors signaling attention from other members of the network (Donath, 2008). For instance, explicit responses to a post via a like or comment is indicative of attention provided to the user. Assessing embeddedness within smartphone interactions is more challenging as there are no explicit markers to signal attention. However, individuals who have a higher social activity can be said to have a greater sense of belonging within the network. Therefore, phone use characteristics such as social activity can be an indicator of social support.

Research also specifies the role played by strong ties in providing social support (Utz & Breuer, 2016). Communicating with strong ties on Facebook is associated with increases in perceived social support and reductions in stress (Burke & Kraut, 2013) This reasoning suggests that online communication with stronger ties rather than with weaker ones should lead to larger improvements in well-being. We find tie-strengths within phone use networks to be an indicator of social support.

A feature of potential interest is the ratio of phone usage during the AM phase (midnight – 11:59 am) to the PM phase. In the considered dataset, the AM phase corresponds to the less socially active time period, and a higher score on this feature corresponds to more than usual activity during this period. Mobile phones allow an individual the possibility of being constantly connected to one's family and friends, however, etiquette dictates that most communications are made in more common or "regular" hours. Therefore, the analysis of phone-use patterns at irregular times could provide useful insights into an individual's social network. In past studies, late night communication has been connected with stress and potentially the need for social support. This suggests a future research direction where the *temporality* of interactions between individuals should be studied for its associations with strength of ties, as well as diversity in network structure.

The second research question for this work (RQ2) focused on the feasibility of a machine learning algorithm to automatically infer support based on phone metadata. Based on the results of the Lasso regression models, we report that phone-based features can perform reasonably well at estimating social support as obtained via survey methods. The correlation between the predicted and the actual overall social support scores was 0.65 and this went up to 0.68 in cases where demography data were also available. As many of the demography variables commonly available to phone service and app providers, can be inferred using phone metadata, or require one-time input, they can often be used in conjunction with phone based features. The resulting model has a MAE of ±0.58 which indicates that the predictions are within ±0.58 of the "ground truth" or survey based values of social support scores

This work is intended to start a conversation on the topic of automated inference of social support, rather than be a final word on it. Hence, we will be cautious in generalizing the results to larger populations until they are verified at scale. We define a set of 17 features for the sample set (n=55) and use Lasso regression, a technique specifically designed to avoid overfitting, and consider the correlation analysis only as a post-hoc interpretation mechanism. We note the moral and ethical considerations in building a social support score based solely on passive data collection. However, this score is intended to help individuals gain an awareness of their social network and the resources they potentially have access to. We also recognize the privacy implications of using an individual's smartphone metadata, however, we suggest using explicit opt-in measures that allow individuals to retain control of their information (e.g. via OpenPDS (de Montjoye, Shmueli, Wang, & Pentland 2014)). While the large-scale analysis and inference of social behaviors requires a larger policy conversation, we believe there is value in exploring these newer directions in order to inform the policy debate.

The phone based predictions obtained here are not an exact replication of the scores obtained by the survey. However, the associations found paint a vision forward and the explanatory power achieved is comparable to those found in similar social computing studies. Given the limitation of sample size, we believe that the correlation scores (0.68) obtained here show a promising direction for the automated prediction of an individual's social support scores.

Despite certain limitations, this paper makes an important contribution towards the literature on social informatics, human centered data-science, and social computing. While previous efforts have studied self-reported interconnections between social behavior and social support, there is no existing effort that has studied the potential of using automated phone-based data to infer and predict the social support scores for individuals.

IMPLICATIONS AND FUTURE DIRECTIONS

In this study we track an individuals' social support via the phone, without having the need for complicated surveys or generator methods, making it a convenient way to get an ongoing assessment of how their daily interactions can impact social support and potentially suggest how it can be increased. The connection between social support and general well-being has been well-documented in literature (Zhang & Yang, 2013; Chang, 2009). It has been shown to directly contribute to the physical and psychological outcomes of health interventions (Zhang & Yang, 2013). The ability to provide and receive social support from a network is also an important motivation for information sharing and social engagement (Oh & Syn, 2013).

For instance, in spite of the many advantages gained from social support, many Americans don't feel they have access to this valuable resource (APA, 2018). When asked if there is someone they can ask for support in times of stress, such as talking over problems or helping make difficult decisions, more than half (55 percent) also said they could have used at least a little more emotional support. Individuals' also reported feelings of loneliness and uncertainty about access to support in stressful situations. American Psychological Association's (2015) article on strengthening supportive networks discusses the importance of proactively building a network that can help provide social support in times of need. Therefore, the ability to track one's social support score via a ubiquitous device like the mobile phone can be useful for individuals to be aware of their social support scores and also give valuable insights on the effect of regular interactions on one's social support score. For instance, people often lose established connections due to life changes such as retirement, relocation, or the death of a loved one. An assessment of how these changes affect one's supportive network can help provide a spur towards forging newer connections in order to have access to a supportive network when required.

Future work in this direction would require a nuanced conversation addressing the constructs of social support as well as integrating more specific phone features such as video calls as well as validation at a larger scale with participants from various demographics. With a larger target group and wide-scale validation, the proposed could help billions of users to keep a track of their social support scores and receive valuable suggestions when appropriate.

REFERENCES

Adler, P.S., & Kwon, S. (2002). Social capital: Prospects for a new concept. *Academy of Management Review*, 27(1), 17–40.

Barrera, M., Jr. (1986). Distinctions between social support concepts, measures, and models. *American Journal of Community Psychology*, 14, 413-416.

Beale, R. (2005). Supporting social interaction with smart phones. *IEEE Pervasive Computing*, 4(2), 35-41.

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073-1076.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood.

Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of computer-mediated communication*, 21(4), 265-281.

Chang, H. J. (2009). Online supportive interactions: Using a network approach to examine communication patterns within a psychosis social support group in Taiwan. *Journal of the Association for Information Science and Technology*, 60(7), 1504-1517.

Chen, J., Hsieh, G., Mahmud, J. U., & Nichols, J. (2014, February). Understanding individuals' personal values from social media word use. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 405-414). ACM.

Cole, D. A., Nick, E. A., Zelkowitz, R. L., Roeder, K. M., & Spinelli, T. (2017). Online social support for young people: Does it recapitulate in-person social support; can it help? *Computers in Human Behavior*, 68, 456-464.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. Psychological bulletin, 98(2), 310.

de Montjoye, Y. A., Quoidbach, J., Robic, F., & Pentland, A. S. (2013, April). Predicting personality using novel mobile phone-based metrics. In *International conference on social computing, behavioral-cultural modeling, and prediction* (pp. 48-55). Springer, Berlin, Heidelberg.

de Montjoye, Y. A., Shmueli, E., Wang, S. S., & Pentland, A. S. (2014). openpds: Protecting the privacy of metadata through safeanswers. *PloS one*, 9(7), e98790.

Donath, J. (2008). Signals in social supernets. *Journal of Computer-Mediated Communication*, 13(1), 231–251.

Dunkel Schetter, C., & Brooks, K. (2009). The nature of social support. *Encyclopaedia of human relationships*, 1565-1570.

Ellison, N. B., Vitak, J., Gray, R., & Lampe, C. (2014). Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. *Journal of Computer‐ Mediated Communication*, 19(4), 855-870.

Ghosh, I., & Singh, V.K. (2016). Predicting Privacy Attitudes Using Phone Metadata. In *Proceedings of 2015 International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction* (SBP15). 51-60.

Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. *Journal of psychosomatic research*, 69(5), 511-520.

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

Haslam, S. A., O'brien, A., Jetten, J., Vormedal, K., & Penna, S. (2005). Taking the strain: Social identity, social support, and the experience of stress. British Journal of Social Psychology, 44(3), 355-370.

Jumin, C., Hyungeun, J., & Jung-hee, R. (2010, February). Predicting Interpersonal Relationships based on Mobile Communication Patters. In *2010 ACM Conference on Computer Supported Cooperative Work (CSCW)* (pp. 487-488).

Kadushin, C. (2002). The motivational foundation of social networks. *Social networks*, 24(1), 77-91.

Kovanen, L., Saramaki, J., & Kaski, K. (2010). *Reciprocity of mobile phone calls*. arXiv preprint arXiv:1002.0763.

Li, H., Ai, W., Liu, X., Tang, J., Huang, G., Feng, F., & Mei, Q. (2016, April). Voting with their feet: Inferring user preferences from app management activities. In *Proceedings of the 25th International Conference on World Wide Web* (pp. 1351-1362). International World Wide Web Conferences Steering Committee.

Lin, N. (1999). Building a network theory of social capital. *Connections*. 22(1), 28-51.

Luarn, P., Kuo, H. C., Chiu, Y. P., & Chang, S. C. (2015). Social support on facebook: the influence of tie strength and gender differences. *International Journal of Electronic Commerce Studies*, 6(1), 37.

Manalel, J. A., & Antonucci, T. C. (2017). Social Relations and Social Support: Understanding Group and Individual Differences. In *Later-Life Social Support and Service Provision in Diverse and Vulnerable Populations* (pp. 28-46). Routledge.

McDowell, I. (2006). *Measuring health: a guide to rating scales and questionnaires*. Oxford university press.

O'Keefe, G. J., & Sulanowski, B. K. (1995). More than just talk: Uses, gratifications, and the telephone. *Journalism & Mass Communication Quarterly*, 72(4), 922-933.

Oh, H. J., Ozkaya, E., & LaRose, R. (2014). How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. *Computers in Human Behavior*, 30, 69-78.

Oh, S., & Syn, S. Y. (2015). Motivations for sharing information and social support in social media: A comparative analysis of Facebook, Twitter, Delicious, YouTube, and Flickr. *Journal of the Association for Information Science and Technology*, 66(10), 2045-2060.

Parks, M. R., & Floyd, K. (1996). Meanings for closeness and intimacy in friendship. *Journal of Social and Personal Relationships*, 13(1), 85-107.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*. 6(1), 65–78.

Quan-Haase, A., Mo, G. Y., & Wellman, B. (2017). Connected seniors: How older adults in East York exchange social support online and offline. *Information, Communication & Society*, 20(7), 967-983.

Ruvalcaba-Romero, N. A., Fernández-Berrocal, P., Salazar-Estrada, J. G., & Gallegos-Guajardo, J. (2017). Positive emotions, self-esteem, interpersonal relationships and social support as mediators between emotional intelligence and life satisfaction. *Journal of Behavior, Health & Social Issues*, 9(1), 1-6.

Salehan, M., & Negahban, A. (2013). Social networking on smartphones: When mobile phones become addictive. *Computers in Human Behavior*, 29(6), 2632-2639.

Stress About Health Insurance Costs Reported by Majority of Americans, APA Stress in America™ Survey Reveals (2018, January). Retrieved from http://www.apa.org/news/press/releases/2018/01/insurance-costs.aspx

Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social science & medicine*, 32(6), 705-714.

Singh, V. K., & Ghosh, I. (2017). Inferring Individual Social Capital Automatically via Phone Logs. (Press) *Proceedings of the ACM Human Computer Interaction*, 1(2).

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* 267-288.

Trepte, S., Dienlin, T., & Reinecke, L. (2015). Influence of social support received in online and offline contexts on satisfaction with social support and satisfaction with life: A longitudinal study. *Media Psychology*, 18(1), 74-105.

Turner, R. J., & Brown, R. L. (2010). Social support and mental health. *A handbook for the study of mental health: Social contexts, theories, and systems*, 2, 200-212.

Utz, S., & Breuer, J. (2016). Informational benefits from social media use for professional purposes: Results from a longitudinal study. Cyberpsychology: *Journal of Psychosocial Research on Cyberspace*, 10(4).

Walther, J. B., & Boyd, S. (2002). Attraction to computer-mediated social support. *Communication technology and society: Audience adoption and uses*, 153188.

Wellman, B. (1992). Which types of ties and networks provide what kinds of social support. *Advances in group processes*, 9(1992), 207-235.

Wellman, B., & Wortley, S. (1990). Different strokes from different folks: Community ties and social support. *American Journal of Sociology*, 96(3), 558-588.

Williams, D. (2006). On and off the 'net: scales for social capital in an online era. *Journal of Computer Mediated Communication*, 11(2), 593–628

Wortman, C. B., & Dunkel-Schetter, C. (1987). Conceptual and methodological issues in the study of social support.

Zhang, M., & Yang, C. C. (2015). Using content and network analysis to understand the social support exchange patterns and user behaviors of an online smoking cessation intervention program. *Journal of the Association for Information Science and Technology*, 66(3), 564-575.

Zhong, E., Tan, B., Mo, K., & Yang, Q. (2013). User demographics prediction based on mobile data. *Pervasive and mobile computing*, 9(6), 823-837.

COPYRIGHT

81st Annual Meeting of the Association for Information Science & Technology | Vancouver, Canada | Nov. 10 - 14, 2018