"Trust Us": Mobile Phone Use Patterns Can Predict Individual Trust Propensity

Ghassan F. Bati
Umm Al-Qura University and Rutgers University
Makkah, Saudi Arabia and New Brunswick, NJ
gfbati@uqu.edu.sa

Vivek K. Singh
Rutgers University
New Brunswick, NJ
v.singh@rutgers.edu

ABSTRACT

An individual's trust propensity - i.e., "a dispositional willingness to rely on others" - mediates multiple socio-technical systems and has implications for their personal, and societal, well-being. Hence, understanding and modeling an individual's trust propensity is important for human-centered computing research. Conventional methods for understanding trust propensities have been surveys and lab experiments. We propose a new approach to model trust propensity based on long-term phone use metadata that aims to complement typical survey approaches with a lower-cost, faster, and scalable alternative. Based on analysis of data from a 10-week field study (mobile phone logs) and "ground truth" survey involving 50 participants, we: (1) identify multiple associations between phone-based social behavior and trust propensity; (2) define a machine learning model that automatically infers a person's trust propensity. The results pave way for understanding trust at a societal scale and have implications for personalized applications in the emerging social internet of things.

Author Keywords
Trust Propensity; Mobile Sensing; Behavioral Sensing

ACM Classification Keywords
J.4 Computer Applications, Social and Behavioral Sciences

INTRODUCTION
Trust is a fundamental human concept that mediates multiple human processes. It facilitates cooperation, supports commerce, and enhances societal well-being [1]. With the growth in social networks, social internet of things, and cyber-physical systems, there is a renewed need to understand and model people's trust propensities as they connect with one another and with the devices around them.

An individual's trust propensity - i.e., "a dispositional willingness to rely on others" - mediates multiple socio-

technical systems [2]. For example, trust propensity strongly influences how an individual makes privacy and security decisions, consumes unverified news, and maintains resources in shared online repositories e.g., [3, 4, 5, 6]. Such scenarios are only likely to grow with the expected growth curves in shared economy, shared augmented reality spaces, and the social internet of things. Hence, understanding and modeling an individual's trust propensity is an important question for human-centered-computing researchers [7, 8].

Multiple recent efforts have attempted to elicit and model an individual's trust propensity using different methods [9, 2]. Nonetheless, such studies have mostly focused on traits which 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 game experiments). Unfortunately, the human-related information taken by observations in such restricted and atypical settings must contend with numerous challenges such as subjective observations, biases, and narrow observation chances while dealing with pressures such as budget, time, and the effort required [10].

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 and cooperation tendencies [11, 12, 13]. In fact, some researchers consider smartphones to be a "vast psychological questionnaire that we are constantly filling out, both consciously and unconsciously" [14].

Given such recent trends and the theoretical literature connecting trust propensity with social capital and social habits such as maintaining interpersonal relationships [15, 16], this study explores the creation of an automated phone data-based approach to model individual trust propensity. Such a phone-based method, if successful, could offer a low-cost, fast, scalable, and automatic method for generating insights into trust propensities for millions of users with applications in social computing, political systems, and sociology.

Hence, building upon this line of work, this exploratory study examines the possibility of using phone metadata to automatically infer an individual's trust propensity by investigating the following research questions:

RQ1: Do long-term phone-use patterns have some associations with an individual's trust propensity?

RQ2: Can a machine learning algorithm be used to automatically infer individual trust propensity based on phone metadata?

In this work, we analyze the data from a ten-week field + lab study to systematically study the interconnections between phone-based behavioral measures (e.g., number of phone calls made) and "ground truth" trust propensity survey scores [17] for 50 individuals.

The rest of the paper is organized as follows. First, we present the related work. Then, we describe the study conducted. Next, we present the obtained results along with their implications and limitations. Finally, we conclude the paper and suggest some potential future work.

RELATED WORK

Trust has been studied across multiple disciplines (e.g., information science, computer science, sociology, psychology, political science, economy) in the past [9, 18, 19, 20, 21]. In this paper, we discuss the related work which is directly connected with the scope of this paper i.e., *modeling trust propensity using phone-based data.* Hence, we discuss the related work that clarifies the terminology and suggests different ways to model trust propensities with a specific focus on computational models of trust. We also review some applications and implications of trust as well as the recent use of mobile phones to infer different behavioral propensities and traits for individuals.

Trust as a Field of Study

Despite its importance and popularity in various disciplines, a clear scientific definition of trust is not obvious [22]. The notions of trust, trust propensity, and trustworthiness are often confused [9, 2, 23]. To remove confusion, we adopt here the following definitions for these concepts:

Trust: "the intention to accept vulnerability to a trustee based on positive expectations of his or her actions" [2, p. 909].

Trust propensity: "a dispositional willingness to rely on others" [2, p. 909].

Trustworthiness: "the willingness of a person B to act favorably towards a person A, when A has placed an implicit or explicit demand or expectation for action on B" [9, p. 65].

While a person's propensity to trust measures their overall willingness to take risks and overall expectations of people to generally behave well, a trustworthy person acts respectfully and with consideration to the needs of other people. In this work, we focus on *trust propensity.*

Trust is an essential social concept for understanding human behaviors in various fields. The presence of trust preserves many relations and produces much good [9]. For example, trust could allow for the use of low-cost informal agreements rather than expensive complex contracts [18]. In addition, individuals in more trusting communities often feel happier and are more content with life, more involved with their local communities, and have more supportive friends [24]. In

computational settings, trust influences purchase patterns in electronic and mobile commerce [25]. Trust is also an important mediator in how individuals deal with security measures, online service agreements, and mobile commerce transactions [26, 27].

Measuring Trust Propensities

Multiple efforts have attempted to elicit an individual's propensity to trust others [2, 9]. However, previous studies have largely focused on demographic traits (e.g., gender, race) or used lab-based experiments (e.g., Dictator Game) [28, 29]. Using such methods for eliciting trust propensity often constrain the scope of studies to factors that can be elicited in the lab settings. Thus, there have been very few attempts that have studied the interconnections between long-term, "in the wild", behavioral features based on mobility or communication traces that range over time and space (e.g., day/night call ratios, average travel distance) and the propensity to trust others.

Computational Modeling of Trust

Several recent efforts have tried to model trust computational settings. Adali et al., define a computational model for interpersonal trust in [7], which treats trust as a social tie between a trustor and a trustee [30]. In this model, trust develops as a part of an emotional relationship between a pair of people akin to the concepts of emotional and relational trust. However, this is quite different from the focus of this paper on *trust propensity,* which is not specific to a relationship, but rather captures an individual's dispositional willingness to rely on (all) others. Similarly, Farrahi & Zia study the *propagation* of trust as a probabilistic stochastic process [31]. Roy et al., propose a pair of complementary measures to determine trust scores of actors in social networks [19] and Zolfaghar & Aghaieb, focus on the evolution of trust in social networks [32]. *However, there are no existing efforts that study the interconnections between individual trust propensities and phone-based data.*

Trust and Social Capital

An individual's trust propensity is often related to their social behavior [15, 16]. A very important concept in the study of social behavior is social capital [33, 34]. Putnam [33] characterizes social capital as trust, network structures, and norms that promote cooperation among actors within a society for their mutual benefit. He also suggests that formal membership, civic participation, social trust, and altruism are indicators of social capital [34]. Such social capital often comes in two variants: bonding and bridging [33]. While bonding social capital is associated with the presence of family and strong personal ties and provides emotional support, bridging social capital is associated with the presence of acquaintances and weak ties that provide access to newer information and resources. Both of these variants of social capital have been connected with trust in multiple studies [33, 35, 36, 37, 38].

Recent HCI studies have connected social capital with phone use behavior suggesting that phone use behavior could also

be predictive of an individual's trust propensity [39, 40, 41]. Trust has also been connected to maintaining inter-personal relationships especially in long distance relationships where face to face interaction is often not possible. Therefore, phone usage patterns could help predict an individual's trust propensity.

Using Mobile Phones to Understand Individual Personality and Propensities

Mobile phones have become a primary communication device used by billions of people globally. Majority of contemporary mobile phones are equipped with several sensors and there exists significant literature utilizing mobile phone sensors to automatically infer individuals' cooperation propensities and personality traits [42, 13, 43]. Indeed, this work builds upon a recent line of work on *phoneotypic* modeling [13] which defines a *phoneotype* as the "*composite of an individual's traits as observable via a mobile phone*". Hence, it argues that a combination of phone-based behavioral features could build a unique signature for an individual which can predict facets of the individual's life (e.g., propensity to cooperate). There are, thus far, no efforts which utilize *phoneotypic,* i.e., phone-based data, to define automated machine-learning approaches for modeling individual trust propensities and this work seeks to address this gap.

STUDY

We study the interconnections between trust propensity and phone-based features based on the data gathered as part of Rutgers Well-being Study undertaken at Rutgers University. This study was a 10-week field and lab study conducted in Spring 2015 including 59 participants, most of whom were undergraduate students at Rutgers.

Initially, all participants were invited to sign consent forms to participate in the study and install an Android app that would record their call, SMS, and GPS logs. Figure 1 shows a screenshot of the app. The app was developed using the "Funf in a box" framework [44] and was released via a URL shared with the study participants.

The participants were also asked to attend three in-person sessions where they filled out a number of surveys concerning their health, well-being, trust propensity, and some demographics. The order of surveys was randomized for the participants. We use here the trust propensity and demographics surveys for their relevance to this work. There was a compensation of US $20, $30, and $50 respectively for attending the sessions.

Participants' privacy was of utmost priority; hence, anonymized IMEI numbers were used to recognize the participants. All user data were anonymized before analysis. Furthermore, the actual phone numbers or the content of the calls or SMS messages were *not* available to the personnel analyzing and processing the data at any point of time. The permissions required for this study's app (call logs, SMS logs, location logs, and phone identifier information) were

intended to be considerably lesser than what is usually required by common apps (e.g., Instagram app on Android). The participation in the study was optional and the participants could withdraw from the study whenever they like. The study was approved by the Institutional Review Board and all personnel who handled the data in this study were trained and certified in human subject research.
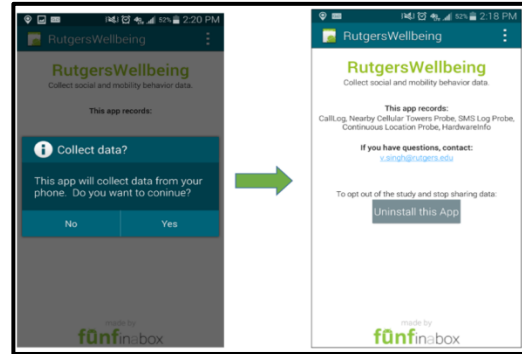

Figure 1. Screenshot of the Android App.

While the study included 59 participants, 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 53 participants. Of these, three participants uploaded location data very rarely (ten or lower instances) - presumably because they turned off location features on their phones - so we removed them from the dataset. This resulted in a dataset involving 50 (32 male, 18 female) 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). Most participants were in the age group of 18 to 21 years, and the most common education level was "some college". The median income of the participants' families ranged from US $50,000 to $74,999.

The 50 participants made a total of 25,302 calls with an average of around 506 and a median of 302.5 calls per participant and exchanged 177,263 SMS messages with an average of 3,545 and a median of 2,347 per participant, and visited 14,045 unique locations with an average of about 280 and a median of 295.5 per participant during the period of the study (10 weeks). Table 1 gives a summary of the total, mean, and median for calls, SMS, and locations.

Feature	Total	Mean	Median
Calls	25,302	506.0	302.5
SMS	177,263	3,545.3	2,347.4
Unique Locations	14,045	280.9	295.5

Table 1. Summary of Calls, SMS, and Location Logs.

Trust Propensity Descriptor

The literature discusses several ways of quantifying an individual's trust propensity. For example, games in controlled lab settings (such as Trust Game) represent one

way of quantifying trust propensities [28, 29]. Surveys that draw individuals' behavior in prepared scenarios are another option [17]. Furthermore, a third way is a combination of game experiments and lab surveys [18].

In this paper, we decided to use a well-known survey "General Trust Scale" to measure trust propensity [17]. The survey has 6 questions whose responses scaled from (5) "Strongly Agree" to (1) "Strongly Disagree" on a five points scale. Some examples of the questions are: "Most people are basically honest" and "Most people are basically good and kind" [17, p. 147]. Besides the prevalent acceptance of the survey (over 1,800 citations as per Google Scholar), we chose this survey as the nature of these questions is not restricted to a specific context and the results could be interpreted in a wide variety of everyday applications. Also, the scale's internal reliability ranges from 0.70 to 0.78 and several studies support its predictive validity [45, 46]. It was developed by selecting items from important trust surveys and has been found to have robust associations with Big Five Personality traits [45, 46].

The scores of the survey are averaged together and normalized as a percentage of the maximum possible score. Thus, the maximum theoretical trust propensity score is 100. In the considered sample, the maximum was found to be 97, the minimum was 40, the mean was 71.5, and the median was 73 as seen in Table 2.

Minimum	Maximum	Mean	Median
40	97	71.5	73

Table 2. Summary of Trust Propensity Scores.

Demographic Descriptors
The participants were surveyed about their demography. Specifically, we obtained the following information: age, gender, marital, level of education (school), and level of family's income.

Mobile Phone Data Features
Trust and socio-mobile behavior have been (indirectly) connected in the past literature in both conceptual and empirical ways. In this work, we consider three major types of socio-mobile features to predict trust propensities.

First, social capital as a concept is connected with both phone use behavior [41] and trust propensities [37, 38]. Hence, we consider a number of phone based features (e.g., number of phone calls, diversity of contacts, and engagement with strong ties) based on the recent literature on using phone meta-data to predict individual social capital or personality traits [43, 13, 41]. In doing so, we do not only consider the frequently used call and SMS metadata, but also consider GPS (location) metadata, which are increasingly being adopted as indicators of physical social activity [47, 48] and also as predictors of an individual's traits and states in their own right [13, 49].

Second, we consider a group of features that have been selected to quantify the trajectories or the mobility behavior of the individuals. These features are related to the concepts of mobility capital (location based analog to social capital) and the notion of a "third place" [50, 51]. (A Third place is a place other than work and home used to build social ties and live a healthy life [51]). Prior research has connected such mobility capital and access to third place with trust [52, 53]. Empirically, these features are based on the recent literature, which has been used to characterize human geo-mobility patterns and study its interconnections with personality and mental health [43, 49].

Third, we consider a set of features that capture the temporal rhythms of human behavior. Conceptually, these features are associated with the notions of circadian rhythms and chronotypes, which have been connected with trust and cooperation in the past literature [54, 55, 56]. Empirically, these features have been based on recent works that have connected similar features with social capital, cooperation, and well-being [41, 13, 57, 58].

We focus on trust propensity which remains largely stable over time. According to [46], trust is an enduring trait, not a transient state. All features here follow a key working assumption based on Macey and Schneider's model connecting states, traits, and behaviors [59]. Propensity (trait) is not transient, but the behavior is affected by both the states and traits. Traits are considered to be long-term predispositions, similar to personality attributes that are often experientially manifested as states, which can be measured indirectly through surveys. States may further manifest through observable and directly measurable behaviors. Hence, we hypothesize that an individual trust propensity traits manifest themselves in the *long-term* behavior patterns of the users [13, 60]. A summary of the features (N=24) is presented in Table 3.

1. Social Behavior

Social Activity
We quantify the level of social activity as the number of exchanged phone calls, SMS messages, and unique visited locations. A higher count of social activity level suggests an active user and multiple studies have connected individual social activity with social capital and/or trust propensities. High social activity has also been connected with reducing relational uncertainty and as a means of establishing trust in interpersonal relationships [22, 61].

We also consider location logs (physical movements) as a proxy of one type of social behavior for it has been used previously to comprehend human social behaviors [48, 49, 13]. The visited locations were updated hourly to balance between getting an idea about the pattern of a user's movement and their phone's battery life.

Type	Literature Support	Features

Social Behavior Features	**Conceptual:** *Social Capital* Putnam [33]; Granovetter [62]; Golbeck [61]; Coleman [63] **Empirical:** Eagle et al. [64]; Shmueli et al. [22]; Gilbert et al. [65]; deMontjoye et al. [43]; Singh & Agarwal [13];	• Social Activity (Call, SMS, GPS)
		   o $\sum$Activity
		
		• Diversity (Call, SMS, GPS)
		   o $D_i = -\sum_j p_{ij} \log_b p$
		
		• Novelty (Call, SMS, GPS)
		   o Percent New Contacts $= \frac{\sum \text{New Contacts}}{\sum \text{All Contacts}} \text{ X } 100$
		
		• Tie Strength (Call, SMS, GPS)
		   o Strong Tie Engagement Ratio $= \frac{\sum \text{communication for highest 1/3 contacts}}{\sum \text{ communication}} \text{ X } 100$
		   o Weak Tie Engagement Ratio $= \frac{\sum \text{communication for lowest 1/3 contacts}}{\sum \text{ communication}} \text{ X } 100$

Spatial Trajectory Features	**Conceptual:** *Mobility Capital* Golbeck [61]; Coleman [63]; *Third Place* Oldenburg [51]; **Empirical:** Pappalardo et al. [66]; Canzian et al. [49]; Singh & Agarwal [13]; Singh et al. [67]	• Gyradius $= \frac{\sum \text{distance from centroid for each location visited}}{\text{number of locations visited}}$ • Percent Long Distance Trips $= \frac{\|\text{Long Distance Trips}\|}{\|\text{All Trips}\|} \text{ X } 100$ • Location Loyalty $= \frac{\sum(\text{time spent in top three locations})}{\sum(\text{time spent in all locations})} \text{ X } 100$ • Percent Time Third Place $= \frac{\sum(\text{time spent in third place})}{\sum(\text{time spent in all locations})} \text{ X } 100$

Temporal Rhythm Features	**Conceptual:** *Circadian Cycles & Chronotypes* Jonassona [54]; Lyons [55]; **Empirical:** Abdullah et al. [58]; Saeb et al. [57]; deMontjoye et al. [43]; Singh & Ghosh [41];	• Diurnal Activity Ratio (Call, SMS, GPS)
		   o $DAR = \frac{\sum \text{Activity when productive}}{\sum \text{Activity when relaxed}}$
		• Weekday/Weekend Activity Ratio (Call, SMS)
		   o $WWAR = \frac{\sum(\text{Call,SMS})\text{in weekdays}}{\sum(\text{Call,SMS})\text{in weekends}}$

Table 3. Summary of *Phoneotypic* (phone-based) Features Defined in this Study.

To avoid getting the same amount of locations per participant (24 locations/day), we only count unique locations. The location data were obtained from a mobile phone's GPS as <latitude, longitude> tuple at fourth decimal point resolution, which roughly corresponds to 10m by 10m blocks [13, 68].

$$\text{Social Activity (Call, SMS, GPS)} = \sum \text{Activity}$$

Diversity
We are not only considering the total amount of calls, SMS messages, and unique locations, but also the diversity (measured as Shannon Entropy) for each one of them, as such a diversity metric has been reported to be associated with multiple personal well-being outcomes and personality traits [12, 69].

$$D_i = -\sum_j p_{ij} \log_b p_{ij}$$

Where p_{ij} is the percentage of social events involving individual 'i' and contact 'j', and 'b' is the total number of such contacts.

Novelty
The growth of networks plays an important role in social capital [70]. Hence, we also consider "new contacts" that are not present in the first four weeks of the data collection period. This feature quantifies how much time users devote to their new contacts as compared to their frequent contacts.

$$\text{Percent New Contacts} = \frac{\sum \text{New Contacts}}{\sum \text{All Contacts}} \text{X} 100$$

Tie Strength
Previous studies have related strength of ties and trust [71]. Such literature underscores the value of maintaining relationships with both strong and weak ties, and each may

yield different types of social capital, and presumably, over periods of time, a propensity to trust others.

Following Williams [72], we connect the concepts of 'bonding' and 'bridging' social capital to those of 'strong' and 'weak' ties as proposed by Granovetter and other researchers [62, 65, 73]. We conjecture that the relative spread (or concentration) of communication with strong (respectively weak) ties may be a predictor of one's propensity to trust others. It is anticipated that a person would devote at least 33% of their time with their top-third most frequent contacts (proxy for strong ties) [13]. Nonetheless, a high score like 85% may indicate an individual's preference to intentionally engage more with strong ties rather than distributing the communication effort more equally amongst all ties. Hence, we define the following features:

$$\text{(STER)} = \frac{\text{Strong Tie Engagement Ratio}}{\sum \text{communication for highest 1/3 contacts}}{\sum \text{communication}} \text{X } 100$$

$$\text{(WTER)} = \frac{\text{Weak Tie Engagement Ratio}}{\sum \text{communication for lowest 1/3 contacts}}{\sum \text{communication}} \text{X } 100$$

2. Spatial Trajectories

Prior research has connected a number of mobility or spatial trajectory related concepts (e.g., mobility capital and access to third place) with trust [52, 53]. Hence, we consider a number of GPS related features to quantify individual behavior.

Gyradius
To get a sense about the location distribution of a participant (physical activity), we determine the gyradius (radius of gyration) which is computed as follows. First, we identify the centroid of all the distinct points that a person has visited. Next, we calculate the distance to all points from this center point. The average of such distances traveled is the gyradius [74].

$$\text{Gyradius} = \frac{\sum \text{distance from centroid for each location}}{\text{number of locations visited}}$$

Percentage Long-distance Trips
An individual's access to new resources and information is likely to be a function of their access to "far-away" people and places. Hence, we also define a feature called Percentage Long-distance trips to quantify the ratio of long distance (above 100 km) trips undertaken by the individual.

$$\text{Percent Longdistance Trips} = \frac{||\text{Long Distance Trips}||}{||\text{All Trips}||} \text{X } 100$$

Location Loyalty
Location loyalty considers how frequently participants engage with their favorite locations. Past research has connected this loyalty feature with individual well-being [75]. Precisely, we calculate the percentage of time spent in their top three frequented visited locations out of all visited locations.

$$\text{LocationLoyalty} = \frac{\sum (\text{time spent in top three locations})}{\sum (\text{time spent in all locations})} \text{X } 100$$

Percentage Time Third Place
We also introduce here the third place feature which represents the percentage of time spent at the third most visited location by a participant. This is based on the sociological concept of "third place", proposed by Ray Oldenburg, which states that a person needs a third place - other than work and home (e.g., library, café, worshipping house) - to build social ties and live a healthy life [51]. Past research has connected third places with social capital and trust [53].

$$\text{Percent Time Third Place} = \frac{\sum (\text{time spent in third place})}{\sum (\text{time spent in all places})} \text{X } 100$$

3. Temporal Rhythms

Prior literature has connected circadian cycles, Dark Triad (i.e., narcissism, machiavellianism, and psychopathy) and trust [54, 55]. The classification of different individual's chronotype - the tendency for the individual to sleep at a particular time during a day-and/or-night period (24-hour) - has been connected with cheating and machiavellianism [56].

Diurnal Activity Ratio
When we asked some of the participants about their daily activities regarding times when they become productive, and times when they tend to play or sleep (relax), we found that there are two main states: "productive" state from 8 am to 8 pm; "relax" state from 8 pm to 8 am. Hence, to quantify daily patterns of activity and the differences between different phases, we define the following features:

$$\frac{\sum (\text{Call, SMS, Location}) \text{when productive}(8am \text{ to } 8pm)}{\sum (\text{Call, SMS, Location}) \text{when relaxed}(8pm \text{ to } 8am)}$$

Weekday/Weekend Activity Ratio
We added another layer of characterization for the abovementioned two states of the daily activity ratio (productive and relaxed) to get more insights out of these circadian rhythms by quantifying the weekdays (Monday to Friday) to weekends (Saturday and Sunday) communication (Call, SMS) ratio.

$$\frac{\sum (\text{Call, SMS}) \text{in weekdays}}{\sum (\text{Call, SMS}) \text{in weekends}}$$

RESULTS
Since multiple applications vary in their requirements of either predicting an exact numeric trust propensity score or working with broader classifications of trust propensity score, we consider both types of applications by undertaking linear regression and classification analyses as follows.

Building a Regression Model for Trust Propensity
Here, we first consider predicting trust propensity level as a regression problem; that is, predicting an outcome variable (i.e., trust propensity level) from a set of input predictors (i.e., phone-based features). We use the Lasso (Least Absolute Shrinkage and Selection Operator) regression approach to

undertake this [76]. Lasso is a specialized form of regression suitable for scenarios where there are relatively more number of features for a given sample size. It tries to minimize overfitting by penalizing the presence of too many features in the eventual model. It has been applied in similar contexts (in terms of sample size, number of features, and application) in recent human-centered computing research [11, 41]. Similarly, following [11, 41] we evaluate the regression models using the metrics of correlation scores (Cor) (between predicted and actual outcome variables) and the Mean Absolute Error (MAE). While a higher correlation (closer to 1) suggests a higher predictive ability of the considered models, smaller MAE is preferred as it shows that the predictions are closer to the ground truth.

We ran and tested three different regression models: one with the demographic features only, another one with the *phoneotypic* (phone-based) features only, and a third one with a combination of both types of features. The implementation was undertaken using R 3.4.1 [77] and its Lars 1.2 package [78]. To test the statistical significance of these three models, we need an estimate of the (variance) in the effects found. To estimate this, we undertook 100-fold bootstrapping for each Lasso regression model and then undertook unpaired t-tests for the correlation and MAE scores obtained. All comparisons were found to be statistically different at alpha= 0.05 level i.e., Both *>* *Phoneotype* *>* Demography (*>* means statistically significantly higher performance). Table 4 presents the average results for modeling trust propensities using various regression models.

The demography based model obtained on average a correlation of 0.274 (MAE=9.146). The low - but significant - scores for the "demography only" model indicates that the demographic features can explain some (but not a lot) of variance in the trust propensity levels. Phone-based model performed much better with an average correlation score of 0.538 (MAE=7.913).

The combined model using *phoneotype* and demography features performed the best in terms of all metrics and the predicted trust propensity was found to have 0.544 correlation on average with the actual propensity scores (MAE=7.711). This MAE signifies that the predictions are within ±7.711 of the absolute value of the trust propensity scores obtained by the survey (ground truth). Since the trust propensity scores obtained by the survey vary from 40 to 97 as shown in Table 2, ranges of ±7.711 could be considered a reasonable approximation.

Also, we clearly see that the *phoneotype* model and "*phoneotype* + demography" (both) models yield considerably better models than the demography-based model. However, the demographic features were useful in increasing the correlation score for the *phoneotypic* model, thus suggesting that *phoneotypic* features and demographic features are not merely proxies for each other, but rather add newer information when combined.

Model Type	Cor	SD	MAE	SD
Demography Only	0.274	0.062	9.146	0.416
***Phoneotype* Only**	0.538	0.153	7.913	1.776
Both	0.544	0.153	7.711	1.541

Table 4. Average Results for Modeling Trust Propensities Using Different Regression Models.

Building a Classification Model for Trust Propensity
Next, we consider the task of building automated classifiers for trust propensities. In prior research, the same Yamagashi trust scale was used to separate participants into groups of high and low trustors [46]. The survey results predicted behavioral differences between groups of individuals. For instance, groups of high trustors were more likely to cooperate and reciprocate across variations of the prisoner's dilemma and public goods problems [46]. This motivates the analysis on the (phone-based) behavioral differences between high and low trustors and create computational models to be used in other applications. For instance, an application provider may want to recommend different default privacy settings for individuals with "high" and "low" trust propensity.

Given that there is no universal definition of "high" and "low" trust propensity, for this exploratory work, we divided the participants into two groups based on the median value (73) for trust propensity survey instrument. The first group ("low" propensity) has 23 participants whose trust score is lower than the median, whereas the second group ("high propensity") has 27 participants whose trust score is higher than or equal to the median.

Demography Only	Age, School (education level)
Phoneotype Only	SMS Entropy, Weekday Weekend Call Ratio
Both	SMS Entropy, Weekday Weekend Call Ratio, Age

Table 5. Selected Features for Different Prediction Models.

Similar to the previous analysis, we built three models: one with the demographic features only, another one with the *phoneotypic* features only, and a third one with a combination of both types of features.

Method	Demography Only			Phoneotype Only			Both		
	AUC	CA	F1	AUC	CA	F1	AUC	CA	F1
AdaBoost	**0.68**	**0.69**	**0.66**	**0.81**	0.76	**0.75**	**0.83**	**0.79**	**0.78**
Random Forest	0.58	0.62	0.60	0.79	**0.77**	**0.75**	0.82	0.78	0.77
KStar	0.63	0.62	0.60	0.73	0.64	0.62	0.80	0.73	0.72
Zero-R	0.50	0.54	0.38	0.50	0.54	0.38	0.50	0.54	0.38

Table 6. Average Results of Trust Propensity Levels Using Different Classification Methods.

We used CfsSubsetEval (Correlation-based Feature Subset Selection) [79] with leave-one-out cross-validation in Weka 3.8.1 [80, 81] which ranks the best subset of the 24 features described previously by determining the predictive capability of each feature in company with the degree of redundancy between them. The best subsets of features are correlated with the target variable and have low intercorrelation [79]. We found that the best subsets of features in most of the folds are the ones shown in Table 5.

To define and test a machine learning based classifier whose *phoneotypic* features can statistically significantly improve the ability of predicting trust propensity when compared to the demographic features, we took 10-fold cross-validation and repeated it 10 times to get 100 different values for CA, AUC, and F1 and build the predictive models. AUC stands for (Area Under the Receiver Operating Characteristic Curve), CA means classification accuracy and F1 score represents the harmonic mean between precision and recall [82, 83]. The aforementioned features were used to test out three well-known machine learning algorithms for classification. Specifically, we used Adaptive Boosting (AdaBoost), Random Forest, and KStar. We also used a Zero-R model which simply classifies all the instances into the majority class, as a baseline to help interpreting the performance of the considered models. Statistical comparison was undertaken using unpaired t-tests (at alpha= 0.05 level) suggesting that for AUC, CA, and F1: *Phoneotype* *>* Demography, Both *>* Demography, Both (not significantly different from) *Phoneotype*; (*>* means statistically significantly higher performance). All three models above were significantly better than Zero-R.

Table 6 shows that the demography-based model returned the best CA of 69%, AUC of 0.68, and F1 score of 0.66. The *phoneotype*-based model yielded a better classification performance and the best CA was 77%, AUC was 0.81, and F1 was 0.75. While the demographic features contained some predictive power, we observe that *phoneotypic* models considerably outperform demographic models.

It is also clear that the *phoneotypic* model outperformed the Zero-R model. The *phoneotypic* model performed 62% better than the Zero-R model in terms of AUC, 42.6% better in terms of CA, and 97.4% better in terms of F1.

We also considered the cases where the demographic data may be available to the phone app. In such a case, the combined model (demography + *phoneotype* data) yielded an even higher performance with a CA of 79%, AUC of 0.83, and F1 of 0.78.

Hence, we note that a phone-features based model beats baseline majority classification and also goes beyond static demographic descriptors (e.g. age, gender, education) for predicting trust propensities. This underscores the potential for using phone-based (*phoneotypic*) features to build automatic classifiers for individual trust propensities. One way to interpret these results is that having mobile sensing data for 10 weeks may allow for the creation of a detailed model for personal behavior based on the aforementioned idea of phone behavior being akin to a vast psychological questionnaire, being constantly filled out [84].

Behavioral Features Associated with Trust Propensity

Besides creating automated methods for identifying an individual's trust propensity levels, one of the goals of this work is to understand the socio-mobile behavior of individuals with different propensities to trust. Thus, we undertook a *post-hoc* Pearson's correlation analysis between trust propensity scores and the *phoneotypic* features. In the interest of space, we only report the correlations that were found to be (at least marginally i.e., p<0.10) significant in Table 7.

Feature	r	p-value
Social Activity (Call)	+0.237^o	0.097
Strong Tie Engagement Ratio (Call)	+0.249^o	0.081
Weekday Weekend Ratio (Call)	+0.371**	0.008
Gyradius	-0.271^o	0.057
Percent Time Third Place	-0.252^o	0.078

Table 7. Pearson's Correlation between Phone-based Features and Trust Propensity. Significance Codes (** 0.01,* 0.05,^{o} 0.10).

We note that people who have high trust propensity tend to be more socially active, yet tend to limit or concentrate their social activities both spatially and temporally. For instance, individuals with higher trust propensity tend to call more often (r= +0.237). This can be understood as trust propensity

is associated with healthy social relationships and higher call activity captures such behavior [24, 18].

Next, we notice that individuals with higher trust propensity tend to have higher preference for *concentrating* social activities in multiple ways. First, they show a marked preference for engaging in phone calls with their "strong ties" as opposed to spending it equitably with all contacts (r= +0.249). The notion of concentrating social activities continues temporally and we notice that the individuals with higher trust propensity tend to concentrate their calling more over the weekdays as opposed to spreading it evenly across all days of the week (r= +0.371).

This aspect of concentrating activities becomes even more prominent when we consider their spatial trajectories. Individuals with higher trust propensity tend to have a smaller gyradius (r= -0.271) and spend less time at even their third-favorite place (r= -0.252), presumably preferring to spend time at their top two favorite locations. One way to interpret these results is that those who travel further and frequently tend to have limited chances to build strong ties and the lack of strong ties has been associated with a lower trust propensity in the past [85].

DISCUSSION

The first research question (RQ1) for this study was: *Do long-term phone-use patterns have some associations with an individual's trust propensity?*

The Pearson's correlation analysis in the preceding section indicates that multiple phone-based features are correlated with an individual's trust propensity. We notice that the individual effect sizes are small and the p-values for multiple of the associations are considered marginally significant. We acknowledge this as a limitation of the sample size, but our confidence is increased by considering that many of the same features show up to be prominent in the features selected by Lasso regression and those selected by the classification algorithms.

Hence, while further testing on individual features is needed as part of the future work, the exploratory work here suggests multiple associations between trust propensity and phone-based social behavior.

The individual associations found can also be connected with the literature connecting trust and social relationships. First, the findings suggest that trust propensity builds more on "strong ties" rather than "weak ties" [62]. While, higher social activity was positively associated with trust propensity, it was also found to grow in concentrated (social, spatial, and temporal) accumulation of such connections. Presumably, repeated social interactions with familiar faces and places, i.e., "bonding" social capital is conducive for developing trust propensities. Conversely, it is possible that those with higher trust propensities tend to build and focus on a small number of relationships.

According to the social identity and self-categorization theories, group-based stereotypes or in-group favoring behaviors might explain how an individual trusts strangers [86]. While individuals normally have good expectations on strangers (out-group members), they anticipate a better treatment when it comes to in-group members (in-group favoritism) which eventually transforms into a greater trust propensity to an in-group, not an out-group member [87, 86, 88]. Constant interactions with such in-group members may result in a longer-term internalization of this trust propensity. All of these aspects are associated with the positive association observed between concentrating social activities - socially and geographically - and a higher trust propensity.

The relational uncertainty theory (RUT), which studies the degree of confidence people have in their perceptions of involvement within interpersonal relationships [89] gives yet another perspective to understand the results. It suggests that trust in long distance relationships is negatively associated with relational uncertainty and reducing uncertainty via constant communication (Social Activity (Call), Strong Tie Engagement Ratio) might be positively associated with trust building. While RUT has mostly been studied in terms of face to face interactions in the past, the current results suggest that similar relationships might hold over phone interactions too.

The second research question (RQ2) for this study was: *Can a machine learning algorithm be used to automatically infer individual trust propensity based on phone metadata?*

The three types of analysis adopted in this work (regression analysis, correlation analysis, and the classification models) suggest that machine learning and in general analytics approaches can indeed be used to infer individual trust propensity based on phone metadata to a large extent. The regression analysis can estimate the individual trust propensities with high correlation (0.544) and within a margin of ±7.711 over a range of 40 to 97. Complementing phone features with demographic data, where available, could yield even better performance. For instance, the classification analysis yielded up to 79% accuracy (AUC=0.83; F1=0.78) based on such models.

Given the modest sample size, we concentrate here on finding general patterns and trends over the three analysis techniques. We can see a consistency in the results across the three analysis methods suggesting that socio-mobile signals as observed via a phone (*phoneotype*) could indeed be used to infer trust propensity of an individual to a reasonable extent.

Privacy of User Data and Ethical Considerations

All data used in this study were hashed and anonymized as discussed in the study design. The permissions needed for the study app were designed to be significantly lesser than those typically adopted by popular apps. Lastly, the participation in

the study was on a voluntary basis and the participants could drop out at any time.

We also note the ethical concerns surrounding assigning an individual a score based on their propensity to trust. While such scores could be used by an individual to receive recommendations for privacy, social networking, and mobile commerce applications, they could also be used by commercial and other organizations to infer individual trust propensities. Similar concerns have been raised about the traditional paper survey based methods administered by any organization, and also newer automated techniques that use social media and phone data to assign health, well-being, or similar "suitability" scores to individuals [90]. Instead of shunning away from reporting such results, or shrouding such research in secrecy, we adopt the approach of raising awareness about these new possibilities and informing the policy debate surrounding them.

Limitations

This study has some limitations. First, we acknowledge that our analysis focused on correlations, not causation. Next, the homogeneity of the sample (participants were mostly undergraduate students from the same university) stops us from generalizing the findings to larger populations, yet it permits isolating socio-mobile behavior as a predictor.

From a methodological perspective, we note the multiple comparisons undertaken in the correlation analysis. While such multiple comparisons are often "corrected" using Bonferroni-Holm correction to maintain the confidence in the associations found, we do not do so here because our analysis is *posthoc* and intended to help interpret the observed prediction results rather than being prescriptive in its own right. Similarly, we acknowledge the issues associated with the use of a relatively large number (24) of possibly collinear features in regression given the modest sample size (50). While this makes the interpretation of individual feature coefficients difficult, the model's average correlation scores of 0.538 for *phoneotype* (respectively 0.544 for *phoneotype* + demography) remain interpretable, especially given the use of Lasso regression, which is purposely designed to handle such scenarios [76].

While we consider the results here to be exploratory, the results from the regression analysis, correlation analysis, and the classification models point to a common theme that there are indeed interconnections between *phoneotype* features and trust propensities. These results motivate further work in this direction.

Implications

With further validation, this line of research could have multiple implications for individuals as well as the society.

We suggest the use of such methods to be based on opt-in. The participants who opt-in to such automated trust propensity scoring apps could get better customized recommendations for privacy, security, social networking, news, and mobile commerce apps. For instance, in [91], the

authors found that trust propensity is an antecedent of the attitudes of mobile users toward in-app ads. Similarly, understanding trust propensity is likely to be the most relevant trust antecedent in contexts involving unfamiliar actors [2]. This is important to understand societal changes and emerging socio-technical contexts like the sharing economy [92]. Generally, the suggested phone-based method here could open ways to better model human beings based on ubiquitous sensing and act as a building block towards the vision of Internet of People [93].

At a societal level, such apps could alleviate the need to run costly annual surveys to access the *trust-based* "state of the nation" as proposed by [18]. Instead, automated methods can be used to create a real-time nation-wide trust propensity census and make it a part of the public policy and decision making process. Further, an ability to study the phenomenon of trust propensity and its "in the wild" dynamics at scale can substantially advance the literature in several fields (e.g., economics, psychology) that study trust propensity. For instance, this approach can help the researchers in many fields to ask research questions that were not simply feasible in labs settings (e.g., contagion in trust propensities across networks of millions of users). In this sense, this work supports the vision painted in the smartphone psychology manifesto, which states that "… smartphones could transform psychology even more profoundly than PCs and brain imaging did" ([84], p.1).

CONCLUSION AND FUTURE WORK

In this work, we have proposed a new approach to infer individual trust propensities using phone features as an alternative to conventional methods like surveys and lab experiments. Using phone-based behavioral features allowed us to build predictive models by means of machine learning classification algorithms whose accuracy, AUC, and F1 scores were promising and encouraging. To the best of our knowledge, there has been no previous study that analyzes the link between individual trust propensity and phone-based behavioral features. Hence, these results pave way for more research on leveraging ubiquitous sensing data for understanding the interconnections between socio-mobile behavioral data and trust propensities.

With further technical and ethical ground work, the proposed approach can be used for inferring trust propensities of individuals at a scale of billions of people. Hence, with the growth in mobile phone penetration, the proposed approach could have multiple implications for individuals (e.g., customized apps) and societies as they engage in higher levels of technology-mediated interactions.

REFERENCES

[1] K. Cook (Ed.), Trust in society, New York, NY: Russell Sage Foundation, 2001.

[2] J. A. Colquitt, B. A. Scott and J. A. LePine, "Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships With Risk Taking and Job Performance," *Journal of applied psychology,* pp. 909-927, 2007.

[3] C. M. Cheung and M. K. Lee, "Understanding consumer trust in Internet shopping: A multidisciplinary approach," *Journal of the American society for Information Science and Technology,* pp. 479-492, 2006.

[4] D. L. McGuinness, H. Zeng, P. P. Da Silva, L. Ding, D. Narayanan and M. Bhaowal, "Investigations into Trust for Collaborative Information Repositories: A Wikipedia Case Study," in *The World Wide Web Conference (WWW2006)*, Edinburgh, UK, 2006.

[5] M. J. Metzger, "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research," *Journal of the American Society for Information Science and Technology,* pp. 2078-2091, 2007.

[6] E. Sillence and P. Briggs, "Trust and engagement in online health a timeline approach," in *The Handbook of the Psychology of Communication Technology*, Oxford, Wiley-Blackwell, 2015, pp. 469-487.

[7] K. Kelton, K. R. Fleischmann and W. A. Wallace, "Trust in digital information," *Journal of the American Society for Information Science and Technology,* pp. 363-374, 2008.

[8] M. Harding, B. Knowles, N. Davies and M. Rouncefield, "HCI, Civic Engagement & Trust," in *the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, New York, NY, USA, 2015.

[9] A. Ben-Ner and F. Halldorsson, "Trusting and trustworthiness: What are they, how to measure them, and what affects them," *Journal of Economic Psychology,* pp. 64-79, 2010.

[10] J. Giles, "Computational social science: Making the links," *Nature,* vol. 488, no. 7412, pp. 448-450, 2012.

[11] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev and A. Campbell, "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, Washington, US, 2014.

[12] Y. A. de Montjoye, S. S. Wang, A. Pentland, D. T. T. Anh and A. Datta, "On the Trusted Use of Large-Scale Personal Data," *IEEE Data Eng. Bull.,* pp. 5-8, 2012.

[13] V. K. Singh and R. R. Agarwal, "Cooperative Phoneotypes: Exploring Phone-based Behavioral Markers of Cooperation," in *The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*, Heidelberg, Germany, 2016.

[14] H. Grassegger and M. Krogerus, "The Data That Turned the World Upside Down," Vice Media, 28 January 2017. [Online]. Available: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win. [Accessed 18 September 2017].

[15] J. Berg, J. Dickhaut and K. McCabe, "Trust, reciprocity, and social history," *Games and economic behavior,* vol. 10, no. 1, pp. 122-142, 1995.

[16] J. Fogel and E. Nehmad, "Internet social network communities: Risk taking, trust, and privacy concerns," *Computers in human behavior,* vol. 25, no. 1, pp. 153-160, 2009.

[17] T. Yamagishi and M. Yamagishi, "Trust and commitment in the United States and Japan," *Motivation and emotion,* pp. 129-166, 1994.

[18] J. Ermisch, D. Gambetta, H. Laurie, T. Siedler and S. C. Noah Uhrig, "Measuring people's trust," *Journal of the Royal Statistical Society: Series A (Statistics in Society),* pp. 749-769, 2009.

[19] A. Roy, C. Sarkar, J. Srivastava and J. Huh, "Trustingness & trustworthiness: A pair of complementary trust measures in a social network," in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, 2016.

[20] D. Gambetta, Trust: Making and Breaking Cooperative Relations, Oxford and Cambridge, MA: Basil Blackwell, 1988.

[21] D. M. Rousseau, S. B. Sitkin, R. S. Burt and C. Camerer, "Not so different after all: A cross-discipline view of trust," *Academy of management review,* pp. 393-404, 1998.

[22] E. Shmueli, V. Singh, B. Lepri and A. Pentland, "Sensing, understanding, and shaping social behavior," *IEEE Transactions on Computational Social Systems,* vol. 1, no. 1, pp. 22-34, 2014.

[23] D. H. McKnight and N. L. Chervany, "Trust and distrust definitions: One bite at a time," in *Trust in Cyber-societies*, Berlin Heidelberg, Springer, 2001, pp. 27-54.

[24] R. I. M. Dunbar, "Breaking Bread: the Functions of Social Eating," *Adaptive Human Behavior and Physiology,* pp. 1-14, 2017.

[25] G. K. Sahi, H. S. Sekhon and T. K. Quareshi, "Role of trusting beliefs in predicting purchase intentions," *International Journal of Retail & Distribution Management,* pp. 860-880, 2016.

[26] C. M. Cheung and M. K. Lee, "Understanding consumer trust in Internet shopping: A multidisciplinary approach," *Journal of the American society for Information Science and Technology,* pp. 479-492, 2006.

[27] C. H. S. Dwyer and K. Passerini, "Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace," in *13th Americas Conference on Information Systems (AMCIS 2007)*, Keystone, CO, 2007.

[28] P. Barclay, "Trustworthiness and competitive altruism can also solve the "tragedy of the commons"," *Evolution and Human Behavior,* vol. 25, no. 4, pp. 209-220, 2004.

[29] F. Exadaktylos, A. M. Espín and P. Brañas-Garza, "Experimental subjects are not different," *Scientific Reports,* vol. 3, 2013.

[30] S. Adali, R. Escriva, M. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. Szymanski, W. Wallace and G. Williams, "Measuring behavioral trust in social networks," in *International Intelligence and Security Informatics (ISI)*, Vancouver, BC, 2010.

[31] K. Farrahi and K. Zia, "Trust reality-mining: evidencing the role of friendship for trust diffusion," *Human-centric Computing and Information Sciences,* pp. 1-16, 2017.

[32] K. Zolfaghar and A. Aghaieb, "Evolution of trust networks in social web applications using supervised learning," in *World Congress on Information Technology 2010 (WCIT 2010)*, Amsterdam, Netherlands, 2010.

[33] R. D. Putnam, "Bowling alone: America's declining social capital," *Journal of democracy,* vol. 6, no. 1, pp. 65-78, 1995.

[34] R. D. Putnam, "Tuning in, tuning out: The strange disappearance of social capital in America," *Political Science and Politics,* vol. 28, no. 4, pp. 664-683 , 1995.

[35] B. Lancee, "The economic returns of immigrants' bonding and bridging social capital: The case of the Netherlands," *International Migration Review,* vol. 44, no. 1, pp. 202-226, 2010.

[36] H. Coffé and B. Geys, "Toward an empirical characterization of bridging and bonding social capital," *Nonprofit and Voluntary Sector Quarterly,* vol. 36, no. 1, pp. 121-139, 2007.

[37] K. Newton, "Trust, social capital, civil society, and democracy," *International Political Science Review,* vol. 22, no. 2, pp. 201-214, 2001.

[38] S. Valenzuela, N. Park and K. F. Kee, "Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation," *Journal of Computer-Mediated Communication,* vol. 14 , no. 4, pp. 875-901, 2009.

[39] S. Yang, S. Kurnia, H. Lee and S. Kim, "Mobile phones and social capital: A preliminary study in south Korea," in *PACIS 2008*, 2008.

[40] Y.-F. Chen, "Social Phenomena of Mobile Phone Use: An Exploratory Study in Taiwanese College Student," in *the Second International Conference on Digital Communication Conference of the National Chung-Cheng University*, Chia-Yi, Taiwan, 2006.

[41] V. Singh and I. Ghosh, "Inferring Individual Social Capital Automatically via Phone Logs," *Proceedings of the ACM Human Computer Interaction,* vol. 1, no. 2, 2017.

[42] W. Z. Khan, Y. Xiang, M. Y. Aalsalem and Q. Arshad, "Mobile Phone Sensing Systems: A Survey," *IEEE COMMUNICATIONS SURVEYS & TUTORIALS,* vol. 15, no. 1, pp. 402 - 427, 2013.

[43] Y. A. de Montjoye, J. Quoidbach, F. Robic and A. S. Pentland, "Predicting personality using novel mobile phone-based metrics," in *6th International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, Washington, DC, 2013.

[44] N. Aharony, W. Pan, C. Ip, I. Khayal and A. Pentland, "Social fMRI: Investigating and shaping social mechanisms in the real world," *Pervasive and Mobile Computing,* vol. 7, no. 6, pp. 643-659, 2011.

[45] A. Montoro, P.-C. Shih, M. Román and A. Martínez-Molina, "Spanish adaptation of Yamagishi General Trust Scale," *Anales de Psicología/Annals of Psychology,* vol. 30, no. 1, pp. 302-307, 2013.

[46] A. M. Evans and W. Revelle, "Survey and behavioral

measurements of interpersonal trust," *Journal of Research in Personality,* pp. 1585-1593, 2008.

[47] S. Scellato, A. Noulas, R. Lambiotte and C. Mascolo, "Socio-spatial properties of online location-based social networks," in *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, Barcelona, Spain, 2011.

[48] G. M. Vazquez-Prokopec, D. Bisanzio, S. Stoddard, V. Paz-Soldan, A. Morrison, J. Elder, J. Ramirez-Paredes, E. Halsey, T. Kochel, T. Scott and U. Kitron, "Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment," *PloS one,* vol. 8, no. 4, 2013.

[49] L. Canzian and M. Musolesi, "Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," in *The 2015 ACM international joint conference on pervasive and ubiquitous computing*, Osaka, Japan, 2015.

[50] V. Kaufmann, M. M. Bergman and D. Joye, "Motility: mobility as capital," *International journal of urban and regional research,* vol. 28, no. 4, pp. 745-756, 2004.

[51] R. Oldenburg, Celebrating the third place: Inspiring stories about the" great good places" at the heart of our communities, Cambridge, MA: Da Capo Press, 2001.

[52] V. Kaufmann and E. D. Widmer, "Motility and family dynamics: Current issues and research agendas," *Journal of Family Research,* vol. 18, no. 1, pp. 111-129, 2006.

[53] R. Oldenburg and D. Brissett, "The third place," *Qualitative sociology,* vol. 5, no. 4, pp. 265-284, 1982.

[54] P. K. Jonasona, A. Jonesb and M. Lyonsb, "Creatures of the night: Chronotypes and the Dark Triad traits," *Personality and Individual Differences,* vol. 55, no. 5, p. 538–541, 2013.

[55] M. Lyons and S. Hughes, "Feeling me, feeling you? Links between the Dark Triad and internal body awareness," *Personality and Individual Differences,* pp. 308-311, 2015.

[56] A. Adan and H. Almirall, "Horne & Östberg morningness-eveningness questionnaire: A reduced scale," *Personality and Individual differences,* vol. 12, no. 3, pp. 241-253, 1991.

[57] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording and D. C. Mohr, " Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study," *Journal of medical Internet research,* vol. 17, no. 7, 2015.

[58] S. Abdullah, M. Matthews, E. L. Murnan, G. Gay and T. Choudhury, "Towards circadian computing: early to bed and early to rise makes some of us unhealthy and sleep deprived," in *the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, Washington, 2014.

[59] W. H. Macey and B. Schneider, "The meaning of employee engagement," *Industrial and organizational Psychology,* vol. 1, no. 1, pp. 3-30, 2008.

[60] N. S. Shami, M. Muller, A. Pal, M. Masli and W. Geyer, "Inferring employee engagement from social media," in *the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Seoul, Korea, 2015.

[61] J. Golbeck (Ed.), Computing with social trust, New York: Springer Science & Business Media, 2008.

[62] M. S. Granovetter, "The Strength of Weak Ties," *American Journal of Sociology,* vol. 78, no. 6, pp. 1360-1380, 1973.

[63] J. S. Coleman, "Social capital in the creation of human capital," *The American Journal of Sociology,* vol. 94, p. 95–120, 1988.

[64] N. Eagle and A. S. Pentland, "Reality mining: sensing complex social systems," *Personal and ubiquitous computing,* vol. 10, no. 4, pp. 255-268, 2006.

[65] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," in *the SIGCHI conference on human factors in computing systems*, Boston, MA, 2009.

[66] L. Pappalardo, F. Simini, S. Rinzivillo, D. Pedreschi, F. Giannotti and A. L. Barabási, "Returners and explorers dichotomy in human mobility," *Nature communications,* vol. 6, 2015.

[67] V. K. Singh, L. Freeman, B. Lepri and A. S. Pentland, "Predicting spending behavior using socio-mobile features," in *IEEE International Conference on Social Computing*, Alexandria, VA, 2013 .

[68] D. Wang, D. Pedreschi, C. Song, F. Giannotti and A.-L. Barabasi, "Human Mobility, Social Ties, and Link Prediction," in *The 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11)*, San Diego, California, 2011.

[69] V. K. Singh, L. Freeman, B. Lepri and A. P. Pentland, "Classifying Spending Behavior using Socio-Mobile Data," *Human Journal,* pp. 99-111, 2013.

[70] N. Lin, Social capital: A theory of social structure and action, New York, NY: Cambridge university press, 2002.

[71] R. Bapna, A. Gupta, S. Rice and A. Sundararajan, "Trust, reciprocity and the strength of social ties: An online social network based field experiment," in *Workshop on Information Systems and Economics*, Shanghai, China, 2011.

[72] D. Williams, "On and off the 'net: Scales for social capital in an online era," *Journal of Computer Mediated Communication,* vol. 11, no. 2, pp. 629-666, 2006.

[73] R. E. Nelson, "The strength of strong ties: Social networks and intergroup conflict in organizations," *Academy of Management Journal,* vol. 32, no. 2, pp. 377-401, 1989.

[74] D. Schulz, S. Bothe and C. Körner, "Human mobility from gsm data - a valid alternative to gps," in *Mobile data challenge 2012 workshop*, 2012.

[75] V. K. Singh, B. Bozkaya and A. Pentland, "Money walks: implicit mobility behavior and financial well-being," *PloS one,* vol. 10, no. 8, 2015.

[76] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 58, no. 1, pp. 267-288, 1996.

[77] R Core Team, *R: A Language and Environment for Statistical Computing,* Vienna, Austria: R Foundation for Statistical Computing, 2017.

[78] T. Hastie and B. Efron, *lars: Least Angle Regression, Lasso and Forward Stagewise,* 2013.

[79] M. Hall, "Correlation-based feature selection for machine learning," April 1999. [Online]. Available: https://www.lri.fr/~pierres/donn%E9es/save/these/articles/lpr-queue/hall99correlationbased.pdf. [Accessed 5 January 2018].

[80] E. Frank, M. A. Hall and I. H. Witten, "The WEKA Workbench," 2016. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf. [Accessed 5 January 2018].

[81] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD explorations newsletter,* vol. 11, no. 1, pp. 10-18, 2009.

[82] Orange Data Mining, "Orange Visual Programming Documentation Reales 3," 31 March 2017. [Online]. Available: https://media.readthedocs.org/pdf/orange-visual-programming/latest/orange-visual-programming.pdf. [Accessed 31 March 2017].

[83] A. Zheng, Evaluating Machine Learning Models A Beginner's Guide to Key Concepts and Pitfalls, Sebastopol, CA: O'Reilly Media, Inc., 2015.

[84] G. Miller, "The smartphone psychology manifesto," *Perspectives on psychological science,* vol. 7, no. 3, pp. 221-237, 2012.

[85] D. Krackhardt, N. Nohria and B. Eccles, "The strength of strong ties," in *Networks in the knowledge economy*, New York, NK, Oxford University Press, 2003, p. 82.

[86] M. Tanis and T. Postmes, "A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour," *European Journal of Social Psychology,* vol. 35, p. 413–424, 2005.

[87] M. J. Platow, M. Foddy, T. Yamagishi, L. Lim and A. Chow, "Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership," *European Journal of Social Psychology,* vol. 42, p. 30–35, 2012.

[88] M. Foddy and R. Dawes, "Group-based trust in social dilemmas," in *New Issues and Paradigms in Research on Social Dilemmas*, New York, USA, Springer Science and Business Media, 2008, pp. 57-85.

[89] L. K. Knobloch and D. H. Solomon, "Information Seeking Beyond Initial Interaction," *Human Communication Research,* vol. 2, no. 28, pp. 243-257, 2002.

[90] Z. Tufekci, "Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency," *Colorado Technology Law Journal,* vol. 13, no. 2, pp. 203-218, 2015.

[91] M. F. Cheunga and W. To, "The influence of the propensity to trust on mobile users' attitudes toward in-app advertisements: An extension of the theory of planned behavior," *Computers in Human Behavior,* vol. 72, pp. 102-111, 2017.

[92] M. Möhlmann and A. Geissinger, "Trust in the Sharing Economy: Platform-Mediated Peer Trust (to be published)," in *Cambridge Handbook on the Law of the Sharing Economy*, Cambrridge, UK, Cambridge University Press, 2018.

[93] J. Miranda, N. Mäkitalo, J. Garcia-Alonso, J. Berrocal, T. Mikkonen, C. Canal and J. M. Murillo, "From the Internet of Things to the Internet of People," *IEEE Internet Computing,* vol. 19, no. 2, pp. 40-47, 2015.