Automatic Classification of Mobile Phone Contacts

Alireza Sahami Shirazi Huy Viet Le, Niels Henze, Albrecht Schmidt

University of Stuttgart, VIS, Stuttgart Germany

Abstract

Current smartphones have virtually unlimited space to store contact information. Users typically have dozens or even hundreds of contacts in their address book. The number of contacts can make it difficult to find particular contacts from the linear list provided by current phones. Grouping contacts ease the retrieval of particular contacts and also enables to share content with specific groups. Previous work, however, shows that users are not willing to manually categorize their contacts. In this paper we investigate the automatic classification of contacts in phones' contact lists, using the user's communication history. Potential contact groups were determined in an online survey with 82 participants. We collected the call and SMS communication history from 20 additional participants. Using the collected data we trained a machine-learning algorithm that correctly classified 59.2% of the contacts. In a pilot study in which we asked participants to review the results of the classifier we found that 73.6% of the reviewed contacts were considered correctly classified. We provide directions to further improve the performance and argue that the current results already enable to ease the manual classification of mobile phone contacts.

1 Introduction

With the increase of mobile phones' ubiquity, they have become one of the most common channels for communication. The wide dissemination of mobile phones also increased the number of contacts that can be assembled in a user's contact list. Contact lists contain the most intimate friends, colleagues, and family members but also other persons that only have a lose connection with the user. Current systems lack an understanding of the relation between a user and her or his contacts. They require that users categorize their contacts manually. Recent mobile phone's operating systems provide tools for manually grouping contact items. However, people do not use these tools. Grob et al. (2009) report that only 16% of users create any contact groups. Min et al. (2013) suspect that users do not perceive enough value to categorize the hundreds of contacts digitally maintained. Further, research shows that relationships are dynamic (Onnela et al. 2007) and need to be periodically updated (Kelly et al 2011). On the other hand, grouping contacts allows users to more easily maintain the privacy and with whom they want to share information. This privacy feature is included in social networks such as Facebook and Google+ but not directly available from phones' contact lists.

In this paper we investigate whether it is possible to automatically group contact items on smartphones based on the communications history (call logs and SMS data) retrieved only from the mobile phone. To define possible groups, we conducted an online questionnaire and asked participants to provide groups they would use to categorize their contacts. Based on the results we select 5 groups (family, friend, acquaintance, work, and other). Then, we recruited 20 participants and collected communication data from their mobile phones. We use machine-learning techniques to classify the contacts with 59.2% accuracy and ask participants of a pilot study to review the results, which reveal an accuracy of 73.6%.

2   Related Work

Researchers have investigated social networks and interaction between their ties. Approaches have focused on tie strength based on four dimensions: amount of time, intimacy, intensity, and reciprocal services (Granovetter 1973). Based on this information distinct groups within a social network are analyzed. Olson et al. (2005) found that people share information with their peers based on the type of relationship. Gilbert and Karahalios (2009) achieved 85% accuracy for a binary classification of weak and strong ties. They suggested that privacy controls based on tie strength might help to divide a user's social network into meaningful groups. Jones and O'Neill (2010) reported six criteria for grouping that people commonly considered: social circles and cliques, tie strength, functional roles, organizational boundaries, temporal episodes, and geographical locations.

Another strand of research evaluates the usability and effectiveness of contact apps. Oulasvirta et al. (2005) and Knittel et al. (2013) assessed the augmentation of contacts with contextual information. Whittaker et al. (2002) conducted interviews and reported that various criteria, such as history of communication, related to selecting important contacts. Komninos and Liarokapis (2009) explored four contextual cues that related to the importance of mobile contacts, i.e., frequency of use, location and temporal context, task and activity context, and personal preferences. Based on a month-long field study, Jung et al. (2008) reported that personalization and reflection of own communication behavior were main factors encouraging users to explore new usage of contacts apps.

Various researchers have used communication logs for grouping contacts. Ankolekar et al. (2009) used the SMS and calls history to develop a social network. Regroup is a system that helps users to create custom groups on-demand in online social networks (Amershi et al. 2012). Different features such as the number of mutual contacts are used for suggesting similar contacts for a group. Maclean et al. (2011) proposed an algorithm for creating social topologies by mining communication history and identifying likely groups based on co-occurrence patterns. Purtell et al. (2011) also followed a similar approach but used emails and photo tags to create the topologies. Roth et al. (2010) used the user's implicit interaction with (group of) contacts for estimating user's affinity. The affinity is used to generate groups. Min et al. (2013) classified the life facet of contacts on a smartphone. With 90% accuracy they could classify contacts with any communication logs. They considered only three life facets, i.e., 'Family', 'Work', 'Social'. Further, they used information retrieved from external resources, i.e., the Facebook social network. In contrast, we derive the contact groups from empirical findings. We consider five groups of contacts. Further, we solely use features ex-

tracted from the mobile phone without any external resources. Additionally, the performance of the classifier was also assessed through a pilot study with potential users.

3   Facets for Contact Grouping

We used an online questionnaire to identify which facets are relevant for grouping mobile phone contacts. Through the questionnaire we aimed to retrieve facets people would want to use to group their contacts. The questionnaire included demographics questions and questions regarding to the frequency of making phone calls and sending SMS. Most importantly we asked participants to list all groups they would like to use in order to group their contacts. The online questionnaire was available for one week (June 12th – June 19th 2012). It was distributed through mailing lists and social networks, i.e., Facebook and Google+. Answering the questionnaire took approximately five minutes.

In total 82 persons (46 female) filled the questionnaire 65% were 18-25 years old and 35% were 26-50 years old. In total 20% of the participants reported to have less than 50 contacts, 29% between 50 and 100 contacts, and 33% reported to have more than 100 contacts in their address book. 78% of the participants sent SMS (several times) daily and 57% daily made calls (several times).

As the responses collected were mainly in German, we translated all answers to English for analyzing the provided contact groups. Furthermore, we harmonized the answers by combining very similar groups (e.g., work, office, job). After harmonizing the answer we found 32 different groups. The five most common groups were: 'Family' (79.3%), 'Friend' (70.7%), 'Acquaintance' (65.5%), 'Work' (56.1%), and 'others' (22%). The identified groups, indeed, overlap with the facets (work, family, social) Oznec and Farnham (2011) proposed for social networks based on qualitative work. We use these five facets for classifying contacts.

4   Data Acquisition

After determining the groups we collected communication logs to develop a classifier that automatically groups contacts based on features derived from the participants' communication history. We developed an application for Android smartphones to collect the required data from participants' mobile phones.

4.1   Apparatus & User study

We developed an Android application to retrieve the communication log and the contacts in mobile phones' address books. The app uses Android's communication history to retrieve the start and end time of all calls as well as the time SMS are sent and the length of the SMS. After retrieving the data the app randomly selects up to 50 contacts the participants at least once communicated with. Further, the app asks the participant to group the contacts and assign each contact to at least one of the five given groups, i.e., family, acquaintance, friend,

work, or other. With this data the ground truth is also gathered. The collected data is sent to a remote server and stored in a central database.

We recruited 20 participants (12 female, average age=22.2, SD=2.30) with an Android smartphone. We asked them to install our app on their phone, complete the procedure, and provide us data. The participants were mainly students or staffs of our university. The procedure took around 30 minutes per participant.

Group	Number of calls	Call duration (second)	Number of SMS	SMS length (characters)
Family	13.9 (23.86)	87.45 (134.22)	3.3 (8.66)	80.57 (56.69)
Friend	16.31 (25.98)	89.80 (150.22)	236.9 (778.02)	73.41 (39.82)
Acquaintance	1.84 (3.53)	42.05 (147.34)	18.3 (77.06)	89.06 (70.68)
Work	3.56 (6.60)	47.78 (86.90)	4.1 (8.62)	80.95 (42.98)
Other	1.95 (2.36)	35.88 (95.30)	0.1 (0.21)	90.0 (67.88)

Table 1: The average number of calls, the average call duration, the average number of SMS and the average length of the SMS in characters. The numbers in brackets shows the standard deviation.
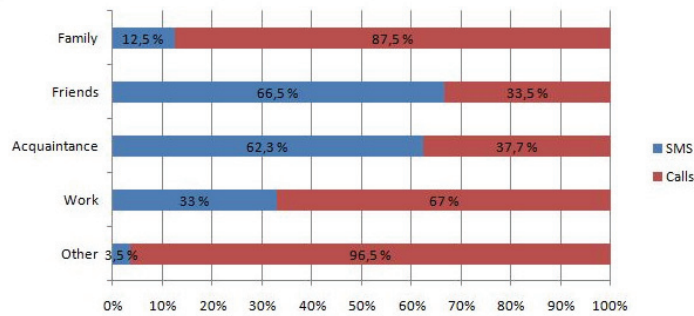

Figure 1: The percentage of SMS and calls in the participants' communication history across the five groups.

4.2   Dataset

The participants in total categorized 439 contacts (M=22, SD=10.49). 153 contacts were classified as 'Acquaintance', 139 as 'Friend', 77 as 'Family', 27 as 'Work', and 43 as 'other'. The collected communication data included 3,895 calls, and 37,189 SMS. An Analysis of Variance (ANOVA) and follow-up post-hoc tests reveal that the number and the duration of calls as well as the SMS' length are significantly different between all five groups (all $p < .05$). There is no significant difference in the number of SMS between the groups ($p > .05$). Table 1 depicts the average number of calls and SMS, the duration of calls, and SMS's length for the five groups. The comparison the ratio between the number of calls and SMS within the group indicates that more than 62% of the communication between contact in the 'Friends' and 'Acquaintance' groups is through SMS. On the other hand for the contacts in the 'Family' group 87% of the communication is through making phone calls. Figure 1 shows the ratio for all five groups.
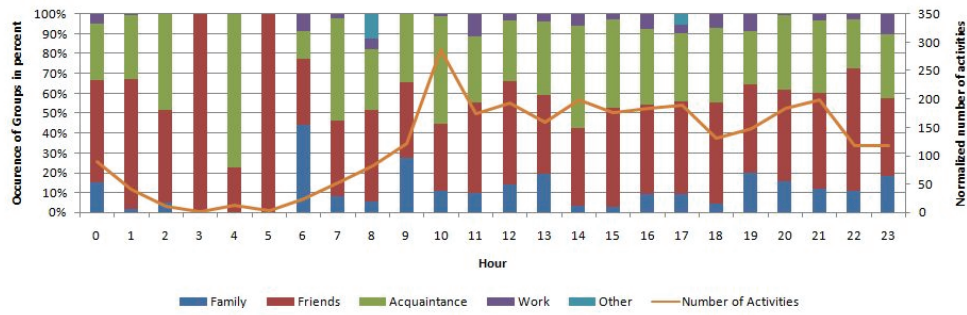
Figure 2: Fraction of SMS that have been exchanged with the five groups over the course of a day.
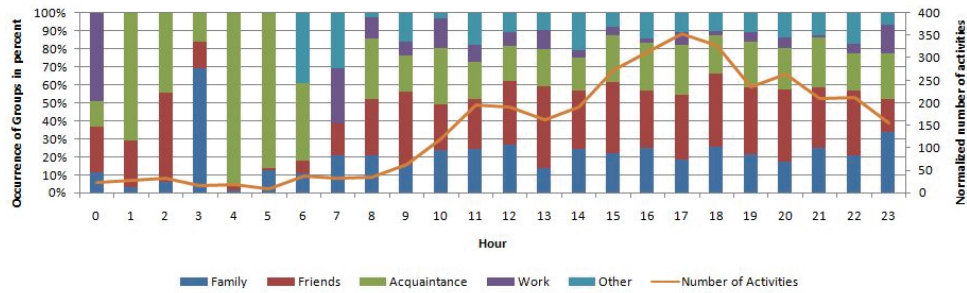

Figure 3: Fraction of calls with contacts from the five groups over the course of a day.

Furthermore, we also assessed the distribution of SMS and call activities during a day. Figures 2 and 3 depict the distribution of sent SMS and calls. The distribution reveals that the participants have more SMS activities in the morning and more call activities in the afternoon and evening. The results show the maximum number of SMS sent and received is around 10 O'clock mainly with Acquaintance, while the maximum call activity is around 17 O'clock with Friends.

Further, we ranked the groups based on the SMS and call activity. To achieve this, we first created a list for each participant and sorted the contacts in a descending order by the number of SMS or call. Then we divided the list into eight equal parts. The first eight contacts got the rank 1, the second eight contacts rank 2, and so on and so forth. Figure 4 depicts the result. The results show that the 'Friend' group is the main contacts in the first and second rank. The contacts in the 'Work' group mainly appear from the rank 3. Interestingly, there is an inverse trend for the 'Family' group in the SMS and call ranking.
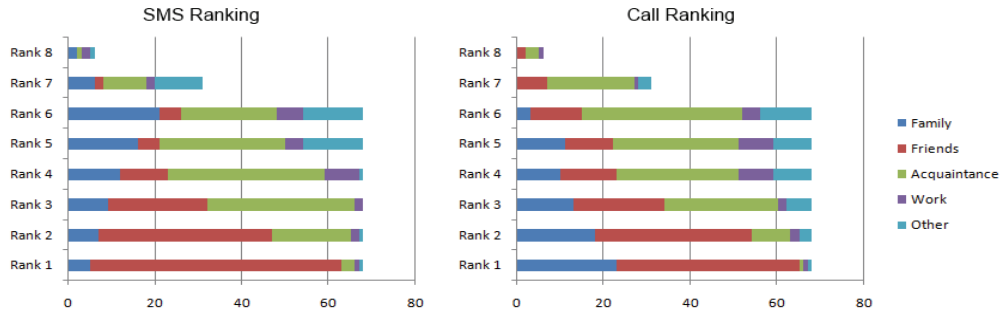
Figure 4: Participants' contacts ranked according to the SMS and call activity.

Feature	Description
Weighted activity	A participant's SMS and call activity in relation to the average SMS activity of all participants.
Ratio between SMS and calls	The ratio between exchanged SMS and calls for a contact in the whole communication history.
Number of calls and SMS	The number of SMS and calls in the communication history of a contact.
Difference between local and average number of SMS and calls	The difference between the number of SMS/calls of a contact and the average number of SMS/calls of the participant.
SMS and call rank	A contact's rank according to SMS and call activity (see previous section and Figure 4).
Average SMS length and call duration	The average length of SMS exchanged with a contact and the average duration of calls with that contact. Unanswered calls are excluded here.
Time	The time of the communication activities is discretized by dividing the day in 4-hour blocks starting at midnight (e.g. 0:00 to 3:59) and determining the block in which most communication with the contact happened.
Weekday or weekend	Described if the communication mainly takes place on weekend or on weekdays by determining when most communication takes place.

Table 3: The features used to train the machine learning algorithms.

5   Classifying Contacts

We used the collected communication logs and the manual classification to derive an automatic classification algorithm. We first derived the features shown in Table 3 from the data. We used the features as input to train a machine-learning algorithm. Experimental results reported in the following are obtained using WEKA (Hall et al. 2009). All learning parameters use the default values in WEKA unless otherwise stated. We used leave-one-out cross-validation to train the classifier and test its performance. That means we trained the classifier

with data from 19 participants and evaluated the performance using the data from the remaining participant. The process was repeated for all participants resulting in 20 runs that were aggregated afterwards.

In a first step we used J48, WEKA's implementation of a C4.5 decision tree, to classify the data. The confusion matrix in Table 4 (left) shows the individual classifications. Overall, 260 contacts (59.2%) were classified correctly and 179 contacts (40.8%) were classified incorrectly. The confusion matrix (see Table 4 left) shows that not all groups can be separated with the same performance. In particular, the groups Family and Others as well as Friends and Acquaintance are difficult to distinguish.

a	b	c	d	e	<- classified as	a	b	c	d	e	<- classified as
44	14	4	5	11	a=family	**31**	18	23	0	6	a=family
7	**116**	9	6	1	b=friends	14	**109**	15	0	1	b=friends
26	24	**81**	10	11	c=acquaintance	9	31	**106**	0	6	c=acquaintance
10	4	6	**4**	2	d=work	9	6	10	**0**	1	d=work
13	1	14	1	**15**	e=other	21	0	13	0	**10**	e=other

Table 4: Confusion matrix resulting from leave-one-out cross-validation using a C4.5 decision tree (left) and a multilayer classifier consisting of three C4.5 decision trees (right).

The classification results suggest that the groups "Friends" and "Acquaintance" as well as the groups "Family" and "Other" share similar characteristics. Users might not be able to clearly separating friends and acquaintance, for example, themselves. Therefore, we constructed a multilayer classifier that consists of three classifiers (see Figure 5). The first classifier (A) only distinguishes between the following three groups: Family & Other, Friends & Acquaintance, and Work. The two additional classifiers take the input from the first one and only distinguish between Family & Other (classifier B) and between Friends & Acquaintance (classifier C). Using the same procedure each classifier has the following accuracy: Classifier A 77.9%, Classifier B 81.7%, and Classifier C 88 %. The resulting classifier correctly classifies 256 out of 439 contacts (58.3%).
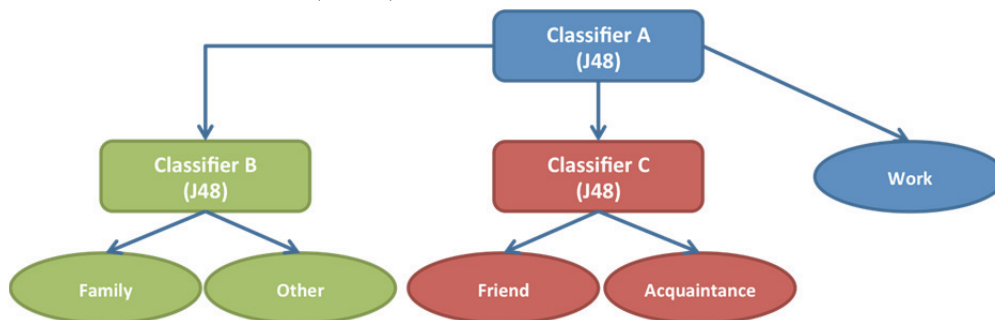

Figure 5: The multilayer classifier consists of three classifiers

6 Pilot Study to Evaluate the Classifier's Performance

It can be assumed that users cannot always precisely categorize their contacts. A user might not necessarily recognize contacts wrongly classified by the classifier as incorrect. To test the performance of the classifier with potential users we conducted a pilot study with five participants. All participants were male students and between 19 and 24 years old.

We implemented an application for Android phones that classifies the contacts using the five groups. The application first collects all contacts with a communication history from the participant's phone and afterwards derives the features. The contacts are classified using the standard C4.5 as well as the multilayer classifier. Then, the participants reviewed the classified contacts. They were asked to correct any result that was wrongly classified.

Overall, the standard C4.5 classifier correctly classifies 93 out of 144 contacts (64.6%) and the multilayer classifier correctly classifies 106 out of 144 contacts (73.6%). The confusion matrixes for both classifiers are shown in table 5.

a	b	c	d	e	<- classified as
12	0	4	0	0	a=family
12	**37**	1	9	4	b=friends
10	0	**38**	2	5	c=acquaintance
0	3	0	**1**	0	d=work
1	0	0	0	**5**	e=other

a	b	c	d	e	<- classified as
10	1	3	2	0	a=family
7	**43**	0	4	0	b=friends
6	0	**47**	6	3	c=acquaintance
1	1	0	**2**	1	d=work
1	1	1	0	**4**	e=other

Table 5: Confusion matrix resulting from participants' review of the output of a C4.5 decision tree (left) and a multilayer classifier consisting of three C4.5 decision trees (right).

7 Conclusion

In this paper we investigated if contacts in mobile phones' address books can be automatically classified using just the communication history consisting of call and SMS logs. Using an online questionnaire we identified the five categories family, friends, acquaintance, work, and other that user would use to categorize their contacts. Collecting real-world data from 20 participants and training machine-learning algorithms we found that 59.2% of the contacts can be correctly classified. Results from a subsequent pilot study in which we asked participants to review the results of the classifier suggest that users are themselves not able to clearly differentiate between the groups. 73.6% of the contacts that have been reviewed by participants have not been recognized as wrongly classified.

Using just the communication history enables to correctly classify the majority of a user's contacts and users would not correct the classification for 73.6% of the contacts. While the performance does not enable a fully automatic process it can clearly lower the burden for the

user. Correcting the category for around a quarter of the contacts is much faster and easier than manually classifying all contacts.

We believe that a (semi-)automatic classification could be clearly improved using individualized classifiers. Users have not only different communication patterns and preferences but might also classify their contacts differently. The classifiers could also be improved by using additional information for training and recognition. Information from social networks such as Google+ and Facebook could help to boost the performance of the classifiers. Furthermore, location information (e.g., derived from area codes) could also help to improve the precision of the classification. Determining contact categories using an online questionnaire might not have determined the exact categories that users would really in daily life. We consider implementing our work in an app that can be deployed on a large scale through mobile application stores (Henze and Pielot 2013, Henze et al. 2013). As in our previous work this could increase the generalizability beyond a specific population (Sahami et al. 2013).

Acknowledgments

References

Amershi, S., Fogarty, J., and Weld, D. (2012). Regroup: interactive machine learning for on-demand group creation in social networks. In Proceedings of CHI.

Ankolekar, A., Szabo, G., Luon, Y. Huberman, B., Wilkinson, D., and Wu, F. (2009). Friendlee: a mobile application for your social life. In Proceedings of MobileHCI.

Gilbert, E. and Karahalios, K. (2009) Predicting tie strength with social media. In Proceedings of CHI.

Granovetter, M.S. (1973). The strength of weak ties. The American Journal of Sociology

Grob, R., Kuhn, M., Wattenhofer, R., and Wirz, M. (2009). Cluestr: Mobile social networking for enhanced group communication. In Proceedings of CSCW.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter.

Henze, N., Sahami Shirazi, A., Schmidt, A., Pielot, M., and Michahelles, F. (2013). Empirical Research through Ubiquitous Data Collection. IEEE Computer, 46(6), 0074-76.

Henze, N., and Pielot, M. (2013). App stores: external validity for mobile HCI. interactions, 20(2), 33-38.

Jones, S. and O'Neill, E. (2010). Feasibility of structural network clustering for group-based privacy control in social networks. In Proceedings of SOUPS.

Jung, Y., Anttila, A., and Blom, J. (2008). Designing for the Evolution of Mobile Contacts Applications. In Proceedings of MobileHCI.

Kelley, P.G., Brewer, R., Mayer, Y., Cranor, L.F., and Sadeh, N.M. (2011). An Investigation into Facebook Friend Grouping. In Proceedings of INTERACT.

Knittel, J., Sahami S. A., Henze, N., and Schmidt, A. (2013). Utilizing contextual information for mobile communication. In Proceedings of CHI EA.

Komninos, A. and Liarokapis D. (2009). The use of mobile contact list applications and a context-oriented framework to support their design. In Proceedings of MobileHCI.

MacLean, D., Hangal, S. Teh, S. Lam, M., and Heer, J. (2011). Groups without tears: mining social topologies from email. In Proceedings of IUI.

Min, J., Wiese, J., Hong, J. and Zimmerman, J. (2013). Mining smartphone data to classify life-facets of social relationships. In Proceedings of CSCW.

Olson, J., Grudin, J., and Horvitz, E. (2005). A study of preferences for sharing and privacy. In Proceedings of CHI.

Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K.,Kertész, J., and Barabási, A.-L. (2007). Structure and tie strengths in mobile communication networks. In Proceedings of the National Academy of Sciences of the United States of America.

Oulasvirta, A., Raento, M., and Tiita, S. (2005). ContextContacts: Re-Designing Smartphone's Contact Book to Support Mobile Awareness and Collaboration. In Proceedings of MobileHCI.

Ozenc, F.K. and Farnham, S.D. (2011). Life "modes" in social media. In Proceedings of CHI.

Purtell, T., MacLean, D., Teh, S., Hangal, S., Lam, M., and Heer, J. (2011). An Algorithm and Analysis of Social Topologies from Email and Photo Tags. In Proceedings of KDD.

Roth, M., Ben-David, A., Deutscher, D., Flysher, G., Horn, I., Leichtberg, A., Leiser, N., Matias, Y., and Merom, R. (2010). Suggesting friends using the implicit social graph. In Proceedings of KDD.

Sahami Shirazi, A., Henze, N., Dingler, T., Kunze, K., and Schmidt, S. (2013). Upright or Sideways? Analysis of Smartphone Postures in the Wild. In Proceedings of MobileHCI.

Whittaker, S., Jones, Q., and Terveen, L. (2002). Contact Management: Identifying Contacts to Support Long-Term Communication. In Proceedings of CSCW.