Friend is Calling: Exploiting Mobile Phone Data to Help Users in Setting their Privacy Preferences

Delphine Christin, Aurélie Bentolila, Matthias Hollick
Secure Mobile Networking Lab
Technische Universität Darmstadt
64293 Darmstadt, Germany
{firstname.lastname}@seemoo.tu-darmstadt.de

ABSTRACT

A continuously increasing number of pictures and videos is being shared in online social networks. Currently, users manually confine access to the contents shared. This configuration process can rapidly become cumbersome for users sharing a large amount of content, and, as a result, they may be tempted to rush through the process or leave the default settings unchanged. This can seriously endanger their privacy if inappropriate users are authorized to access sensitive data. In order to reduce the burden on the users as well as enhance their privacy protection, we propose to leverage information already available on their mobile phone as a basis for recommendations on how to set their privacy preferences. To this end, we conducted a user study exploring the differences between users belonging to different social groups, in terms of communication patterns. We have designed classifiers based on mobile phone data to distinguish members of different social groups, and we have evaluated these classifiers using a real-world dataset. The results show that friends can be easily identified using call and short messages logs, while identifying colleagues requires additional information.

1. INTRODUCTION

In recent years, the public interest for online social media has continuously increased and has led to an unprecedented amount of content being generated and shared by users. Picture and video sharing has become particularly popular, as shown by the estimated 135,800 pictures uploaded every minute to Facebook [1] and the approximated 48 hours of video shared on YouTube every minute [3]. In existing sharing platforms, users protect their privacy by selecting who is able to access each uploaded content element. For example, Facebook users can decide to share pictures with specific individuals, friends, friends of friends, or everyone. The definition of access control rules for the uploaded content can, however, rapidly become time-consuming and cumbersome for the users as the amount of uploaded content increases. Users may therefore be tempted to keep the default settings or to arbitrarily select their privacy settings in order to shorten the selection process. Such behavior can seriously endanger the privacy of the users if the selected settings allow inappropriate persons to access sensitive content.

In order to reduce the overhead for the users and enhance their privacy protection, we propose to support the users in the customization of their privacy settings. We aim at suggesting personalized settings based on information already available in the mobile phones of users, such as phone call logs. This information enables us to form groups of users, who interact according to the same patterns, therefore potentially belonging to the same social group. For example, by using our approach, users could select a person to share a picture with and obtain a list of persons showing the same interaction patterns, similar to friend suggestions on Facebook. Users would only have to select suggested persons in the proposed list, instead of searching for each person individually. Within the scope of this work, we therefore explore the feasibility of identifying the nature of the relationships between users based on their interaction pattern. To this end, we analyze the interaction patterns of the 119 participants of our user study with their family, friends, and colleagues in order to identify characteristics particular to each social group. We build interaction profiles for each social group based on the identified characteristics and we validate them against a real-world dataset. The results show that call and short message interactions enable the identification of friends with good precision, while identifying colleagues appears to be more difficult.

The remaining of this paper is structured as follows. We analyze the design space in Section 2 and present the modalities and results of our empirical user study in Section 3. We detail our validation methods and results in Section 4 and discuss open issues in Section 5. After summarizing existing work in Section 6, we make concluding remarks in Section 7.

2. DESIGN SPACE ANALYSIS

Multiple devices and settings can be used to detect interactions between users, e.g., static deployments of specialized devices. Among the multiple possibilities, we believe that mobile phones are devices well-suited to monitor interactions between users because of their unobtrusiveness and ubiquity. Recent mobile phones can also be leveraged to present the privacy setting suggestions to the users when they upload content to share from their mobile phone to the Internet. Additionally, mobile phones contain a gold mine of information about the social interactions of the users. Even simple mobile phones record call information, such as the name of the caller or the callee, the call time and the call duration. Similarly, the storage of short messages reveals the name of the sender/receiver, the time and the content of each message. Advanced phones can also detect nearby users and provide more fine-grained information about the users based on their email traffic and activities in online social network applications. Each detail of these different interactions can provide hints about the nature of the relationships between users. The degree of granularity of the hints, however, depends on the type of information collected. For example, the content of emails and messages can reveal abundant detail about user relationships through analysis of the lexicon used. However, this analysis requires complex processing due to the difference in vocabulary between users and the possible utilization of different languages while using the same phone. On the other hand, simple hints, such as the frequency of short messages or duration of calls, may be sufficient to classify users in different categories. We decided to first examine whether primitive information about meetings, calls, or short messages are sufficient and reliable enough to distinguish categories of users. We envision extending the scope of this study to more complex information in future work.

3. PRELIMINARY USER STUDY

The goal of this study is to identify relevant characteristics of the interaction between users in order to classify them into three different social categories, namely friends, colleagues, and family. In particular, we have focused on the pattern of their meetings, calls, and short messages, respectively. We recruited the participants of this study in our university and a partner university in France by posting announcements on multiple forums and mailing lists. 119 participants answered our paper/online questionnaire anonymously. The majority of the participants were male (77%) and students (71%). The remaining participants were either employees (23%) or PhD students (6%). All participants use at least one mobile phone, while 12% use two phones. More than 95% of the participants use their mobile phone(s) daily, while the rest uses mobile phones on an irregular basis. Note that we approached this sample of participants specially, since they are potential users of online sharing applications and may benefit from privacy settings suggestions. In the following sections, we examine the answers of the participants and analyze which characteristics of interactions between users, such as frequency, temporal distribution, duration, and location of the interactions, provide sufficient insight needed to classify users into the three proposed categories. Based on these results, we summarize our findings in the form of an interaction profile, which helps to categorize users into the different social categories. Note that the participants answered the survey based on their phone logs and/or memory. While diary studies or a direct analysis of their phone logs would have provided results with a higher degree of accuracy, our primary objective was to identify general trends for the design of interaction profiles, which did not require highly accurate information at this stage.

3.1 Meeting Patterns

In the first step, we analyzed the frequency and the temporal distribution of meetings between family members, friends, colleagues. Note that we refer to *colleagues* as *fellow students* for the participants of our study who are students. We excluded missing and unspecified answers for the following analysis. Confirming our expectations, 58% of the participants indicated that they mostly meet their colleagues on a daily basis, as shown in Figure 1. They also tend to meet their friends daily (38%) or weekly (44%). The daily meeting of friends at workplaces or universities makes the distinction between colleagues and friends difficult when only

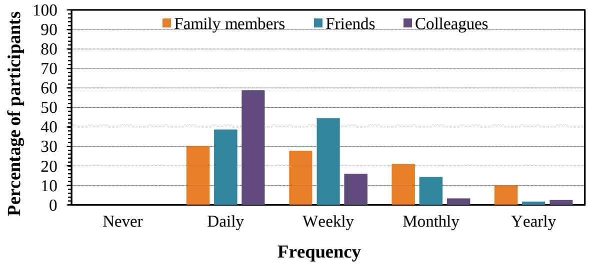

Figure 1: Frequency of meetings in function of the nature of the social relationship

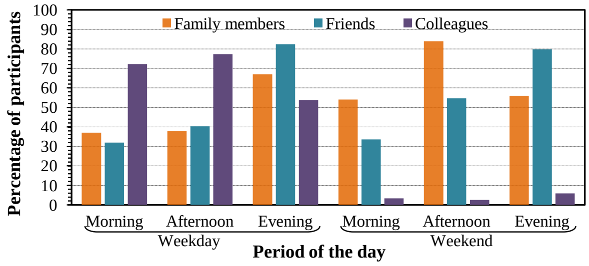

Figure 2: Temporal distribution of meetings (multiple choices possible)

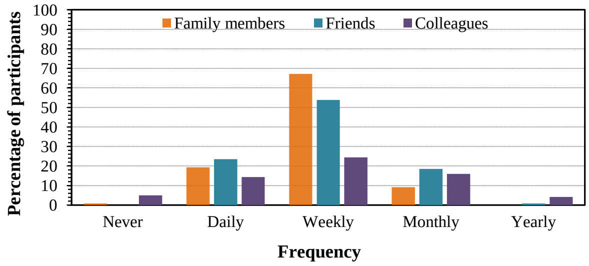

Figure 3: Frequency of phone calls

considering the frequency of meetings. Based on the answers of the participants, it is also difficult to distinguish family members from friends. Figure 2 confirms that a majority of participants meet their colleagues during weekdays and, in particular, in the morning (72%) and afternoon (77%), which obviously corresponds to common working/studying hours. However, 53% of the participants meet their colleagues in the evening as well. This may be explained by extra working hours or shared recreational activities, such as after-work/student parties. The answers of the participants show that they meet their friends mostly in the evenings of both on weekdays (82%) and during weekend evenings (80%), while they meet their family in the afternoon during weekends (84%) and in the evening on weekdays (67%).

3.2 Phone Call Patterns

Figure 3 shows the frequency at which the participants indicated calling members of their family, friends, and colleagues. Similarly to the results presented in Figure 1, the difference between the three categories remains limited. It may, however, be noticed that participants tend to call their family (67%) and friends (53%) on a weekly basis. Figure 4

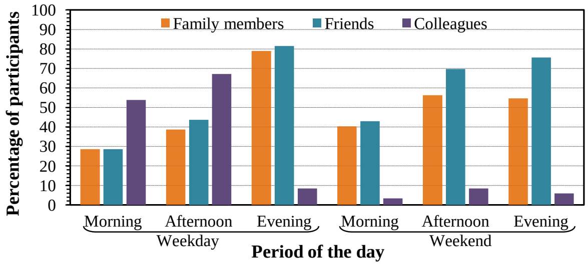

Figure 4: Temporal distribution of the calls (multiple choice possible)

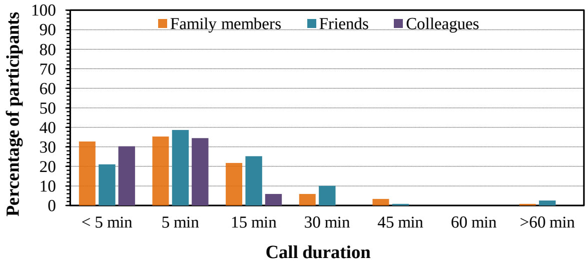

Figure 5: Estimated average call duration

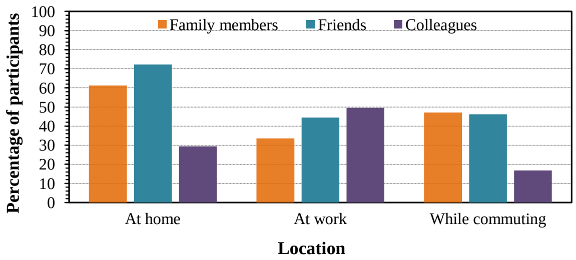

Figure 6: Location while calling (multiple choice possible)

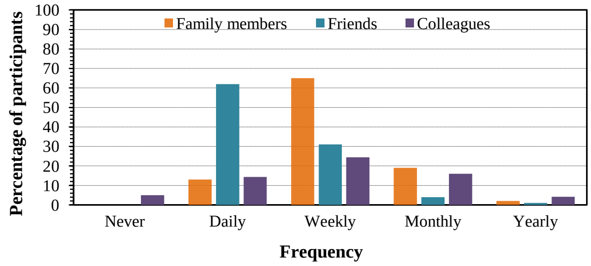

Figure 7: Frequency of short messages

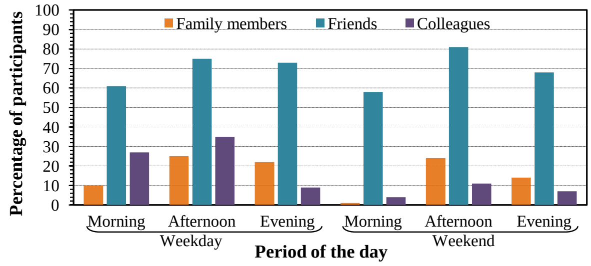

Figure 8: Temporal distribution of short messages (multiple choice possible)

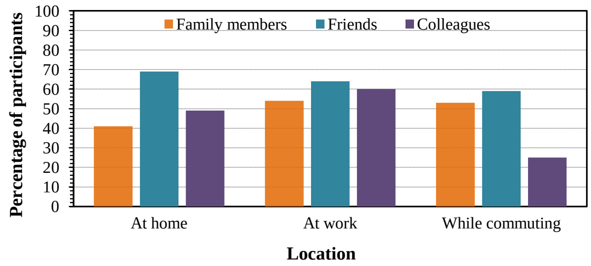

Figure 9: Location while sending short messages (multiple choice possible)

illustrates the temporal distribution of the calls of the participants during a week. The participants call their colleagues mostly during weekdays (53% in the morning, 67% in the afternoon), while they call their family and friends mostly in the evenings and on the weekend. Calls to friends and family follow the same pattern, making the distinction between members of these categories difficult. Moreover, Figure 5 shows only few differences between the duration of the calls between the three categories, except that calls to colleagues do not exceed 30 minutes, according to the answers of our participants. We have also analyzed whether the location of the participants while calling may reveal insights about the recipient of their call. Figure 6 shows that the participants prefer to call their family and friends while at home or commuting, and their colleagues while at work. They still, however, call friends and family while at work. Again, the colleagues are easier to distinguish from both friends and family than friends from family members.

3.3 Short Message Patterns

We next examined the characteristics of the short message pattern. Figure 7 shows the frequency of short message exchange between family, friends, and colleagues, respectively. The answers of the participants illustrate that most participants (62%) exchange short messages with their friends daily, while exchanges with family are mainly weekly (65%). This difference in frequency may therefore be a good indicator for distinguishing friends from family members. This trend is confirmed in Figure 8, which shows their temporal distribution within a week and clearly indicates that the participants primarily send short messages to their friends as compared to their family and colleagues. Figure 9 illustrates the location of the participants when they write short messages. There are slight differences in the location from which the participants send short messages, depending on the nature of their relationships. It seems that they prefer sending short messages to their friends (59%) and family (52%) rather than to their colleagues (25%) while commuting, and to their friends once at home (69%). The differences at work between all categories remain, however, limited.

Based on the results of the user study, we have identified different characteristics of the interaction patterns, which are summarized in Table 1. These characteristics allow us to

Table 1: Profile summary

		Family	Friends	Colleagues
Meetings	Frequency	-	Weekly/Daily	Daily
	Period of the day	Weekday evenings Weekend afternoons	Weekday evenings Weekend evenings	Weekdays
Calls	Frequency	Weekly	Weekly	-
	Period of the day	Weekday evenings Weekend	Weekday evenings Weekend	Weekday mornings Weekday afternoons
	Duration	-	-	< 30 min
	Location	At home While commuting	At home While commuting	At work
Messages	Frequency	Weekly	Daily	-
	Period of the day	Afternoons Weekday evenings	Weekdays Weekends	Weekday afternoons Weekday mornings
	Location	At work While commuting	Everywhere	At work At home

derive different interaction profiles between the participants and their family, friends, and colleagues, respectively.

4. CLASSIFICATION

We have leveraged the results of the aforementioned user study to design classifiers aimed at distinguishing users who have social relationships of varying natures. We have also evaluated the performances of the designed classifiers based on a real-world dataset. In this section, we first provide details about the utilized dataset and the design of the classifiers, and then present the outcomes of their evaluation.

4.1 Dataset

For our evaluation, we used the Reality Mining dataset from the MIT Media Lab [2, 4]. This dataset is composed of fine-grained and multi-modal information about social interactions between real users. It includes the locations of these users, their call and short messages logs, and the IDs of nearby persons. Moreover, each participating user indicated whether he was friend or colleague with other users he interacted with. No information is provided about family relationships, though, and we therefore focus on the distinction between friends, colleagues, and others. In total, the dataset contains 96 active users, including 28 pairs of friends and 83 pairs of colleagues.

4.2 Classifier Design

The results of the user study presented in Section 3 have shown that the participants of our user study interact with their friends and colleagues according to different patterns. For example, they exchange more frequently short messages with their friends than their colleagues. We have analyzed the differences in terms of interaction and translated them into classifiers. Note that similar interaction profiles may have been inferred from the Reality Mining dataset. The existing dataset does not, however, contain information about family members. Moreover, we aimed at evaluating the interaction profiles derived from our user study against a real-world dataset.

The primary goal of a classifier is to distinguish members that have different social relationships, e.g., friends and colleagues, friends and others, or colleagues and others. Each classifier refers to at least one variable (e.g., the total number of meetings of the user (see Table 2)) and defines a condition for each variable (e.g., the number of meetings should be greater than the average number of meeting per user). The designed classifiers are categorized according to their respective interaction modality. Table 2 illustrates the classifiers based on meetings, while Tables 3 and 4 illustrate those based on calls and short messages, respectively. The first column of each table represents the variable(s) of interest. The second column corresponds to the associated condition(s). For each variable, we have chosen different set of conditions, which are defined based on the mean and standard deviation of the corresponding variable (noted m and σ, respectively). Based on the evaluation results of the aforementioned classifiers, we have designed additional classifiers, which combine variables related to multiple interaction modalities. Table 5 shows call- and message-based classifiers, while Table 6 shows classifiers based on all interaction modalities, i.e., meetings, calls, and short messages.

4.3 Evaluation Results

We have applied the designed classifiers on the chosen dataset and evaluated their performances by measuring the true positive, false positive, true negative, and false negative rates. A higher true positive rate and true negative rate corresponds to a better performance by the classifier. Table 2 shows that the meeting-based classifiers provide poor results for both the identification of friends and colleagues. In fact, classifiers 1 to 4 correctly recognize only 1% of existing friends, while classifiers 5 and 6 do not recognize any friends. Classifier 7 presents the best results with only 5% of true positives, though. The classifiers, however, tend to better recognize colleagues than friends, with up to 7% of true positives for classifier 4.

In comparison, call-based classifiers presented in Table 3 provide better results than meeting-based classifiers. For example, classifier 18 based on the number of calls per month shows the best results with 70% of true positives and 90% of true negatives for the identification of friends. Classifiers based on the number of calls on weekdays and weekends (i.e., classifiers 20 to 22) and based on the number of calls in the afternoon/morning and in the evening of weekend days (i.e., classifiers 23 and 25) also perform well with more than 50% of true positives. The same classifiers, however, do not perform well to recognize colleagues. For example, classifier 23 presents 63% of true positives for friends, but only 6% of true positives for colleagues. Classifier 19 presents the best results for the identification of colleagues in this category of

Table 2: Meeting-based classifiers (TP: percentage of true positive, FP: percentage of false positive, TN: percentage of true negative, FN: percentage of false negative)

Classifier number	Variable	Condition	Friends				Colleagues			
			TP	FP	TN	FN	TP	FP	TN	FN
1	x: total number of meetings	x>m	1	99	99	1	2	98	94	6
2		x<m	1	99	99	1	6	94	98	2
3		x- $\frac{\sigma}{2}$<x<m	1	99	99	1	4	96	94	6
4	x: number of weekday morning meetings y: number of weekday afternoon meetings	m- σ <x<m m- σ <y<m	1	99	99	1	7	93	95	5
5		m<x<m+ σ m<y<m+ σ	0	100	99	1	6	94	95	5
6		x>m y>m	0	100	99	1	6	94	95	5
7	x: number of weekend evening meetings y: number of weekend afternoon meetings z: number of weekday evening meetings	m<x m<y m<z	5	95	99	1	0	100	95	5

Table 3: Call-based classifiers

Classifier number	Variable	Condition	Friends				Colleagues			
			TP	FP	TN	FN	TP	FP	TN	FN
8	x: average call duration	x>m	20	80	88	12	4	96	90	10
9		m<x<m+ $\frac{\sigma}{2}$	32	68	88	12	5	95	92	8
10		m- $\frac{\sigma}{2}$<x<m	12	88	84	16	14	86	96	4
11		m<x<m+ σ	32	68	90	10	3	97	91	9
12		x>m+ σ	4	96	83	17	4	96	92	8
13	x: average weekend call duration	x>m	33	67	90	10	0	100	90	10
14		m<x<m+ $\frac{\sigma}{2}$	50	50	90	10	0	100	92	8
15		m<x<m+ σ	43	57	91	9	0	100	91	9
16		x>m+ σ	0	100	84	16	0	100	92	8
17	x: number of calls per month	x>m	39	61	92	8	11	89	93	7
18		x>m+ σ	70	30	90	10	0	100	92	8
19		m<x<m+ $\frac{\sigma}{2}$	9	91	85	15	27	73	94	6
20	x: number of weekday calls y: number of weekend calls	x>m y>m	52	48	92	8	5	95	92	8
21		m<x<m+ σ m<y<m+ σ	50	50	87	13	13	88	93	7
22		x>m+ σ	67	33	89	11	0	100	92	8
23	x: number of weekend morning/afternoon calls y: number of weekend evening calls	x>m y>m	63	38	92	8	6	94	8	92
24		m<x<m+ σ y=0	10	90	85	15	30	70	94	6
25		x>m+ σ	54	46	89	11	0	100	92	8

classifiers with 27% of true positives. This low performance results from the design of the classifiers, which are tailored to primarily identify friends based on the friend interaction pattern highlighted in the aforementioned user study.

Table 4 shows that most of the classifiers based on the exchange of short messages perform better than the previous classifiers. Except for classifiers 31 and 33, the classifiers succeed in identifying at least 75% of friends included in the dataset. Classifier 29 (based on the number of short messages sent during the weekend) and classifier 34 (based on both the number of short messages sent during the week and the weekend) show the best results with 100% of true positives and 76% and 79% of true negatives for friends, respectively. This result confirms the first observations made in the user study, which has highlighted intensive exchange of short messages between friends (see Section 3.3). Both classifiers, however, failed in identifying colleagues as shown by their true positive rate equal to 0%.

Based on these results, we have attempted to improve the performance of the classifiers by combining different interaction modalities in one classifier. Table 5 shows the classifiers, which combine both the call and short message modalities. All combined classifiers present better results than those based on one interaction modality only. They identify 100% of existing friends with at least 98% of true negatives. Finally, we examined classifiers that combine the three interaction modalities, as illustrated in Table 6. The main design driver for these classifiers has been to improve classifier performance when identifying colleagues. In fact, the classifiers based on a unique interaction modality have shown particularly poor results in recognizing this social group. Classifiers 39 and 40 of Table 6 show the best results for the identification of colleagues by reaching 56% of true positives and 97% of true negatives. The combination of the three interaction modalities not only allows the recognition of colleagues, but also the recognition of friends with a good accuracy, as shown by classifier 41 (TP=100%, TN=98%).

Table 4: Short Message-based classifiers

Classifier number	Variable	Condition	Friends TP	Friends FP	Friends TN	Friends FN	Colleagues TP	Colleagues FP	Colleagues TN	Colleagues FN
26	x: total number of sms	$x>m$	88	12	86	14	0	100	86	14
27		$m+\sigma<x$	86	14	82	18	0	100	86	14
28	x: number of sms sent during the weekend	$x>m$	86	14	82	18	14	86	91	9
29		$m+\sigma<x$	100	0	76	24	0	100	88	12
30		$m<x<m+\sigma$	75	25	72	28	25	75	92	8
31	x: number of sms sent on weekdays	$x>m$	57	43	87	13	7	93	87	13
32		$m+\sigma<x$	80	20	75	25	0	100	88	12
33		$m<x<m+\sigma$	44	56	70	30	11	89	90	10
34	x: number of sms sent during the weekend y: number of sms sent on weekdays	$x>m$ $y>m$	100	0	79	21	0	100	88	12

Table 5: Call and message-based classifiers

Classifier number	Variable	Condition	Friends TP	Friends FP	Friends TN	Friends FN	Colleagues TP	Colleagues FP	Colleagues TN	Colleagues FN
35	x: total number of sms y: mean weekend call duration z: number of calls during the weekend u: number of sms sent during the weekend v: number of calls per month	$x>m$ $y>m$	100	0	99	1	0	100	96	4
36		$x>m$ $z>m$	100	0	98	2	0	100	96	4
37		$u>m$ $z>m$	100	0	98	2	0	100	96	4
38		$u>m$ $v>m$	100	0	98	2	0	100	96	4

Table 6: Meeting, call, and message-based classifiers

Classifier number	Variable	Condition	Friends TP	Friends FP	Friends TN	Friends FN	Colleagues TP	Colleagues FP	Colleagues TN	Colleagues FN
39	x: number of sms sent during the weekend y: number of weekday calls z: number of weekday evening calls u: number of weekday afternoon meetings v: number of weekday morning meetings w: mean meeting duration a: number of sms sent on weekdays b: number of weekday evening meetings	$x=0$ $0<y<m$ $z>0$ $m<u<m+\sigma$ $m<v<m+\sigma$ $m-\frac{\sigma}{2}<w<m$	11	89	98	2	56	44	97	3
40		$x=0$ $0<y<m$ $z>0$ $m<u<m+\frac{\sigma}{2}$ $m<v<m+\frac{\sigma}{2}$ $m-\frac{\sigma}{2}<w<m$	11	89	98	2	56	44	97	3
41		$x>1$ $a>0$ $m<u<m+\sigma$ $m<b<m+\sigma$	100	0	98	2	0	100	96	4

In summary, the analysis of a unique interaction modality is not sufficient to identify both friends and colleagues. Either call-based or message-based classifiers provide good results in the identification of friends. Their results can still be improved by combining two interaction modalities. The utilization of two different classifiers, each of them combining the three interaction modalities, has shown encouraging results for the identification of both friends and colleagues.

5. DISCUSSIONS & OPEN ISSUES

In the above evaluation, we compared the performance of 38 different classifiers designed to distinguish users belonging to different social groups. Our goal was to explore different categories of classifiers to assess their respective performances. We plan to refine the design of the classifiers that have shown good results in this exploratory evaluation, in order further improve their performance. Additionally, we plan to include classifiers specially tailored for the recognition of family members. Such classifiers have not been addressed in this work since the available dataset did not contain the relevant information. To the best of our knowledge, no dataset with the same degree of granularity as the Reality Mining dataset exists while also providing information about all three social relationships. This means that evaluating classifiers for the three social relationships under the same conditions would require repetition of the experiment conducted by the MIT Media Lab over several months while considering not only interaction with friends and colleagues, but also with family members.

Creating a new dataset would cater for precious insights into the design of classifiers, but would not be sufficient to reach our ultimate objectives. While the classifiers allow the identification of trends in the interaction patterns of different social relationships, the interaction patterns remain personal to each user, making the development of universally applicable classifiers difficult. These classifiers should therefore be complemented by machine learning algorithms directly running on the mobile phones of users and taking their feedback into consideration. The classifiers could serve as initialization parameters for the machine learning algorithms in order to speed up the learning process, which may appear cumbersome to the users. The learning process is, however, necessary in order to enhance the accuracy of the proposed classification. Moreover, it may ultimately reduce the overhead associated with the customization of the privacy settings, once the algorithms are trained. The introduction of machine learning algorithms opens the doors to novel challenges. This includes their design and implementation, but also the design of usable interfaces for the users to provide feedback, the analysis of the acceptance of the proposed concept by potential users, and the measurement of resulting overheads for the users compared to existing solutions as conducted in [9].

Assuming trained classifiers were able to distinguish friends from colleagues, an additional challenge could be to refine the classification based on the degree of privacy required for each colleague and friend. In this case, a fine-grained analysis of the type of information shared with each individual, e.g., the content of short messages or emails, would be necessary to infer the strength of relationships within the same social group. Moreover, a further challenge would be to determine which information to share with friends, family, and colleagues. Depending on the context, users may be willing to share private information with their family and friends, or with their colleagues. Again, this may be inferred by monitoring and analyzing the information exchanged by each user in order not to only suggest users, but also information to share with other users.

6. RELATED WORK

Existing work can be divided into two main categories. The first category focuses on identifying and/or analyzing friendship relationships based on information collected using mobile phones. For example, *FriendSensing* [8] makes recommendation about potential friends based on proximity information, e.g., frequency and duration of meetings. In comparison, [5] leverages proximity information, location, calls, and short messages to categorize social relationships into three categories: non-friendship, asymmetric friendship, or symmetric friendship. [7] builds on the same categories, but exclusively utilizes data related to explicit communication between users, such as short messages and calls. All of these works do not, however, intend to help users in setting their privacy preferences. In contrast, the second category of existing work concentrates on this aspect. For example, [6] and [9] help users customize their privacy settings based on a set of rules defined by the users themselves. Both solutions, however, require a direct interaction of the users for the definition of the rules. In summary, this work is, to the best of our knowledge, the first one to propose an approach where information available on the mobile phone is leveraged to help users to personalize their privacy settings.

7. CONCLUSIONS

In this paper, we have explored the feasibility of helping users in personalizing their privacy preferences based on data recorded by mobile phones. By means of a user study, we have highlighted the differences in the interaction patterns between members of distinct social groups. We have leveraged these differences to design classifiers, which aim to identify users from different social groups, and we evaluated the performance of these classifiers using a real-world dataset. The results of using a combination of meeting, call, and short message information are promising. In the future, we plan to integrate additional interaction modalities, such as email communication, in order to refine the profile of each person built by the classifiers.

8. ACKNOWLEDGMENTS

Our thanks go to the participants of the user study and the anonymous reviewers for their valuable feedback. This work was supported by CASED (`www.cased.de`).

9. REFERENCES

[1] A Snapshot of Facebook in 2010. Online: `http://www.facebook.com` (accessed in 03.2012).

[2] The Reality Mining Project. Online: `http://reality.media.mit.edu` (accessed in 03.2012).

[3] YouTube Statistics. Online: `http://www.youtube.com/t/press_statistics` (accessed in 03.2012).

[4] N. Eagle and A. Pentland. Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[5] N. Eagle, A. Pentland, and D. Lazer. Inferring Friendship Network Structure by using Mobile Phone Data. *Proc. of the National Academy of Sciences*, 106(36):15274–15278, 2009.

[6] L. Fang and K. LeFevre. Privacy Wizards for Social Networking Sites. In *Proceedings of the International Conference on World Wide Web (WWW)*, pages 351–360, 2010.

[7] S. Mirisaee, S. Noorzadeh, A. Sami, and R. Sameni. Mining Friendship from Cell-Phone Switch Data. In *Proceedings of the International Conference on Human-Centric Computing (HumanCom)*, pages 1–5, 2010.

[8] D. Quercia and L. Capra. FriendSensing: Recommending Friends using Mobile Phones. In *Proceedings of the Conference on Recommender Systems (RecSys)*, pages 273–276, 2009.

[9] R. Ravichandran, M. Benisch, P. G. Kelley, and N. M. Sadeh. Capturing Social Networking Privacy Preferences. In *Proceedings of the International Symposium on Privacy Enhancing Technologies (PETS)*, pages 1–18, 2009.