"Tell Me What I Want to Know!": The Effect of Relationship Closeness on the Relevance of Profile Attributes

João Guerreiro, Daniel Gonçalves Technical University of Lisbon / INESC-ID Rua Alves Redol, 9. 1000-029, Lisboa, Portugal. joao.p.guerreiro@ist.utl.pt, daniel.goncalves@inesc-id.pt

ABSTRACT

The growing amount of personal information on the web raises increasing concerns about what and with whom we share information online. Nevertheless, little effort has been made in determining the relevance of the information shared with us or in filtering it accordingly. We conducted a study to identify the most relevant characteristics when seeking information about people and to scrutinize their differences among relationship types. To achieve that, we asked users to describe people (friends, acquaintances and famous people). Afterwards, we asked them to rate the perceived relevance of a carefully pre-determined set of attributes for each type. Results showed that their relevance varied depending on the relationship. As an outcome, we present the most relevant attributes when seeking information about friends, acquaintances and famous people and the major differences among them. We conclude suggesting how our findings may influence the design of interactive systems where such data is paramount.

Categories and Subject Descriptors

H5.m [Information Interfaces and Presentation (e.g., HCI)]: Miscellaneous.

General Terms

Design, Human Factors, Theory.

Keywords

Personal Information, Social Relationships, Profile Attributes, Social Networking Sites, Relationship Closeness, Person Search, User Study, Information Overload.

1. INTRODUCTION

The web, and particularly Social Networking Sites (SNSs), comprise a large amount of our Personal Information (PI). This fact prompted increasing privacy concerns among users and researchers [1]. Users do not want to freely disclose their PI, reserving their selves the right to choose with whom they share it [10]. The development of mechanisms to specify with whom we share our information (e.g. *Google+ circles* and *Facebook lists*)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

24th ACM Conference on Hypertext and Social Media 1–3 May 2013, Paris, France. Copyright 2013 ACM 978-1-4503-1967-6/13/05 ... \$15.00 enabled sharing content with appropriate audiences within different areas [7], even though many users end up sharing with all their contacts [5].

This large amount of PI, much of which is publicly available, supports the frequent need for information about people. Several search engines (e.g. $pipl^1$) and research projects (e.g. [2] [11]) were created in order to find and compile information about people. Moreover, previous research reported that traditional search engines are regularly used to search for people [8]. Indeed, it also occurs in SNSs, as Facebook users rely on it, regularly, to know more information about their contacts [4].

Although users are now able to manage their information outflow, little headway has been done to help them manage its inflow. With circles and lists, users are able to select the group of contacts they want to browse; still, they cannot specify the type of information they are interested in. Previous research acknowledged that profile attributes differ in relevance, either to predict the number of friends in SNS [6] or to form impressions about people [9]. Although these are valuable insights, it is not possible to generalize such results to scenarios where we seek information about people. Thus, what is lacking is the evidence on which attributes to consider in these conditions. Narrowing the search space with potentially useful information, would decrease the users' effort to find the relevant one. In fact, it would allow warning users (or highlighting) when some contact's important attribute changes. To cite a couple examples, it may occur that in the midst of so much information, we may miss that a good old friend got married or an ex-colleague got a new job.

In a previous analysis, we identified the most relevant attributes when seeking information about friends [3]. However, with the number of contacts always rising, the accessible information goes beyond the barrier of our closest relationships. We believe that the information that we find relevant about a particular individual depends on the type and closeness of the relation we have with that person. We acknowledge that relevance of attributes is not exclusively dictated by the relationship closeness; oppositely, we believe that a wider spectrum of attributes can subsequently be narrowed by other elements, such as context.

In this paper, we present a study aimed at finding and comparing the most relevant attributes when enquiring about people with different intimacy levels (friend, acquaintance and famous). Participants were asked to describe people and to rate the relevance of a set of attributes for a person search scenario. Their combination allowed us to obtain, for different relationship types, the attributes used when describing someone and the ones found relevant when seeking person related information. A comparative

¹ http://pipl.com (Last Visited in 02/2013)

analysis demonstrated that the type of data we find relevant depends on the relation we maintain with that particular person.

2. "TELL ME WHAT I WANT TO KNOW"

We conducted a study that aimed to find out which are the most relevant attributes that we intend to see when searching for information about people. The knowledge and intimacy levels differ among relationships, fact that we took into consideration in order to check if those differences influenced the type of information that we want to know. The study provides the following contributions:

- The most relevant characteristics when seeking information about a friend, acquaintance and famous person;
- The differences among relationship types;

2.1 Methodology

We recruited 23 volunteers (14M/9F), with ages between 19 and 58 (m=32; std.dev=12.2) and different backgrounds (yet, half were Computer Sciences students). They were asked to describe different people, enabling us to find which features are usually mentioned when talking about someone. Moreover, they filled a questionnaire, where we collected additional information about the relevance of several pre-determined attributes from the participants' perspective. In what follows, we describe our methodology; for a more detailed explanation check our previous paper exclusively focused on friends' attribute relevance [3].

2.1.1 Interviews – Describing People

The interviews aimed to lay out the attributes that people usually refer to when describing someone and find out eventual differences between descriptions of people with distinct relationship types. We asked participants to "describe a close friend, providing both profile and previous interactions information". This was carefully worded to try to elicit as much information as possible. Participants also had to describe an acquaintance and a famous person.

2.1.2 Friend, Acquaintance and Famous Person

Every user was asked to describe three different persons. The main reason to describe a *Friend*, an *Acquaintance* and a *Famous person* was to identify the differences arising when talking about people we might have different knowledge levels (from intimate to generally-available public information). We wanted to analyze if the characteristics mentioned when characterizing people with different kinds of relationship differ. By selecting these three types, we focus only on people of which the user has some previous knowledge by two main reasons: first and most obvious, it would be impossible to ask participants to describe someone they do not know; second, we believe that the absence of knowledge/relation boosts the context importance regarding which the most relevant items are.

Participants could choose the friend and acquaintance they described, but the famous person had to be selected from a list of seven. All of them are well known by most people and belong to different areas such as cinema, music, sports, television or politics so they could choose the one that best suits them. *Robert de Niro*, *Madonna* and *President Barack Obama* are three of the persons in this list. We opted to restrict the choice in order to have some degree of control and get realistic descriptions of "regular" famous persons instead of sometimes fantasized and over-hyped descriptions of some idol or the season's media darling.

2.1.3 Pre-Determined Characteristics

We selected a set of pre-determined features which helped us to better catalog the descriptions' data, easing the analysis and allowing to reach to information that participants knew but did not promptly mention. We based our selection in areas related to people descriptions, such as person ontologies (e.g. FOAF), SN profiles (e.g. *Facebook*) and previous studies on profile attributes [15] [22]. The features are presented in Figure 1.

2.1.4 Spontaneous Versus Induced

We classified the information that was revealed in two main categories: *Spontaneous*, when it an attribute was mentioned without any intervention from the interviewer, and *Induced*, if it directly resulted from some question or remark on his part. The interviewer had a set of pre-determined questions that could be asked to try to elicit forgotten relevant features (e.g. "What do you know about his past/origins?"). This setting allowed us to perceive the reasons for not mentioning the attributes spontaneously (e.g. forgetfulness, lack of knowledge). Moreover, the attributes could be classified as *Not known Spontaneous* (mention only to clarify they do not know it) or *Not Mentioned*.

2.1.5 Relevance Questionnaire

This questionnaire intended to know what participants find relevant when they need information about a close friend, an acquaintance or a famous person. This complements the previous task where we wanted to know what they themselves use to describe those people. Both sets are not necessarily the same, and it would be interesting to analyze the eventual differences.

We asked them to imagine they needed information about a close friend, an acquaintance and a famous person, and to classify the relevance of each of the pre-determined characteristics, using a 5-point *Likert Scale*, ranging from 5-*Very Relevant* to 1-*Irrelevant*. This step allowed us to identify attributes that people find relevant but for some reason did not mention on their descriptions. Moreover, it allowed us to validate the relevance of the attributes that were spontaneously mentioned in the interview.

3. RESULTS

Herein we present our study's results. First, we present the attributes that stood out for each relationship type. Then, we specify the attributes with the most evident differences among relationship types.

3.1 Friends

Figure 1 (left) shows the number of participants that mentioned each attribute (left) spontaneously (SPT) or induced (IND) and the median of questionnaire's ratings (right). For a detailed analysis of friends' results check [3]. All in all, by analyzing both interviews and questionnaires, the most relevant features when seeking information about friends are: Personality, Relationship, Interests & Hobbies, Academic History, Profession, Phone, Email and Address. On the other hand, Religion, Political Ideology, Physical Description and Past Addresses were not mentioned many times and were rated as indifferent.

3.2 Acquaintances

Acquaintances (Figure 1 - center) have a similar distribution of SPT and IND in comparison with Friends. Herein, the most mentioned attributes were Profession, Relationship, Personality, Age and Related People. Also promptly remembered by more than half participants, but with less IND are Academic History, Professional Area and Family Information. Other features such as

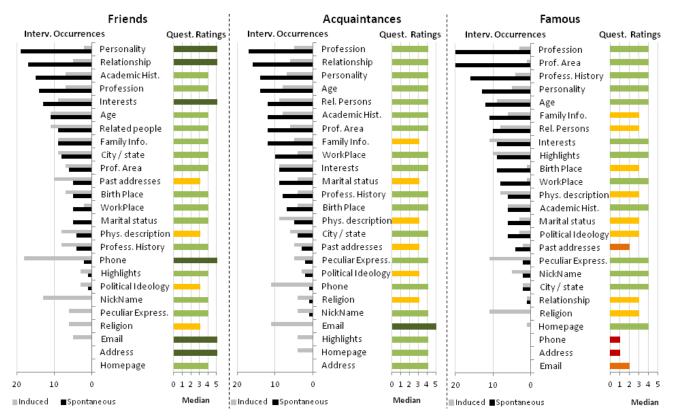


Figure 1. Spontaneous and Induced frequencies for each characteristic (left) and the median classification of their relevance (right), regarding Friends, Acquaintances and Famous People

Workplace, Interests and Professional History were also frequently recalled, especially Interests that was also mentioned many times by induction. Address, Homepage, Email and Highlights did not have SPT occurrences, while the others had just a few. Phone and Email stood out by the high IND value.

The questionnaire confirmed these two features relevance, mainly *Email*, that was the only classified as *Very Relevant* (MD=5) by more than half of the participants. Most other attributes were classified as *Relevant* (MD=4). Those considered *Nor Relevant or Irrelevant* (MD=3) were: *Religion, Political Ideology, Past Addresses, Physical Description, Marital Status* and *Family Information*. The last two had reasonable scores in the interviews but were classified as indifferent, showing they knew that information, but did not find them relevant for *Acquaintances*.

To sum up, the most relevant features when seeking for information about an *Acquaintance* are: *Profession*, *Relationship*, *Personality*, *Age*, *Related People*, *Academic History*, *Professional Area* and *Email*. Apart from *Religion*, *Political Ideology*, *Past Addresses*, *Physical Description*, *Marital Status* and *Familiar Information* (rated indifferent), all other features are also relevant.

3.3 Famous

When describing famous people, the participants' biggest concern was professional (Figure 1 - right), as almost everyone mentioned *Profession, Professional Area* and *Professional History*. Also spontaneously mentioned by more than half participants were *Personality* and *Age*. Others such as *Family Information, Related People, Interests* and *Highlights* were also promptly mentioned. The communication features (*Email, Address* and *Phone*) were never mentioned since participants had no way to communicate with them. *Homepage* and *Religion* were only referred by

induction and *Relationship*, *City*, *Nickname*, *Peculiar Expressions* and *Past Addresses* were mentioned by only a few users.

The questionnaire ratings suggest that the set of relevant features for *Famous* is considerably smaller than for *Friends* and *Acquaintances*. While the most mentioned attributes (professional ones, *Age* and *Personality*) were classified *Relevant* (MD=4), others well positioned were found *Nor Relevant or Irrelevant* (MD=3), such as *Family Information*, *Related People* and *Birth Place*. With lower ratings are *Phone*, *Address* (MD=1), *Email* and *Past Addresses* (MD=2). Among other relevant features are: *Interests*, *Highlights*, *Workplace* and *Academic History*.

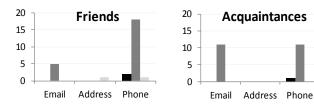
Overall, when seeking information about famous people the most relevant attributes are: *Profession*, *Professional Area*, *Professional History*, *Personality* and *Age*. Other relevant ones are: *Interests*, *Highlights*, *Workplace and Academic History*. Alike with the other relationship types, attributes such as *Nickname*, *Peculiar Expressions* and *Homepage* were not frequently mentioned, probably due to the absence of that information, but were considered relevant.

3.4 Differences among Relationship Types

In what follows, we present the main differences regarding the attributes relevance among relationship types.

3.4.1 Communication Patterns

E-mail, *Address* or *Phone* were not mentioned many times spontaneously (Figure 2). The participants mentioned them by induction mainly when questioned about how they (would) communicate. We will focus on *Friends* and *Acquaintances*, due to the (understandable) poor results regarding the *Famous* type.



■ Spontaneous ■ Induced Not Know Spontaneous Figure 2. Number and type of Communication Features occurrences for Friends and Acquaintances.

Phone was prevalent when talking about friends, whilst for acquaintances Email and Phone are balanced. This suggests that people are more comfortable in using the Phone (a faster and more intimate interaction method) with Friends while for Acquaintances they are divided between Phone and Email, a more distant method. The questionnaire supports that Phone is considered more relevant for Friends (md=5) than Acquaintances (md=4). A Friedman Test (applied also with Famous results) and post-hoc Wilcoxon Signed Rank Tests revealed a significant difference for Phone relevance between Friend and Acquaintance with a large effect size (p<.001, r=.52). The same tests applied to Email did not reveal significant differences between Friend and Acquaintance (md=5). Address ratings were statistically different with a medium effect size (p<.001, r=.48) between Friend (md=5) and Acquaintance (md=4). In what follows, whenever we refer to statistical significance about the questionnaire, these were the tests applied.

3.4.2 Relationship between User – Person Described Concerning the participant-person Relationship, it is coherent the lack of mentions on Famous type. However, two participants talked about some similarities between them (e.g. "he has the same disease I have"). The number of references was the same for Friends and Acquaintances (22), but with more SPT for the former (17 against 16), emphasized in the questionnaire where users rated it more relevant for Friends (md=5) than Acquaintances (md=4). The questionnaire differences were statistically significant with medium effect size (p<.05, r=.41).

3.4.3 Interests & Hobbies

Although with a slight advantage, the *Interests & Hobbies* prevalence goes to *Friends* (mentioned by 22) and a probable reason is that in 52% of the times the user shared at least one interest/hobby with the friend described (against 44% and 35% to *Famous* and *Acquaintances*, respectively). There were more *IND* responses than *SPT* for *Famous*, suggesting they knew that information but did not find it relevant. This assumption is supported by the questionnaire in which participants classified it as *Very Relevant* for *Friends* (md=5) and *Relevant* for *Acquaintances* and *Famous* (md=4). These differences are significant with large effect size for *Friend-Famous* (p<.001, r=0.53) and with medium effect size for *Acquaintance-Famous* (p<.01, r=.39) and *Friend-Acquaintance* (r=.41).

3.4.4 Professional History and Area

Professional History and Area had very similar results. They were two of the most unaided recalled attributes for Famous with 16 and 20 SPT mentions respectively, but reduced a lot when talking about Friends (4 and 8) and Acquaintances (6 and 12). We conducted a McNemar-Bowker's Symmetry Test to identify differences between the different relationship types regarding the possible categories (SPT, IND, Spontaneous not Known and Not

Mentioned). It showed significant differences (p<.05) between Famous and Friend and suggests a minor significance between Famous and Acquaintances (p=.053). In what follows, when claiming significance among relationship types in the interviews, this is the test we applied. The questionnaire did not found disparities for Profession History and Area (md=4), as they were rated as Relevant for all, but with a slight advantage to Famous.

3.4.5 Academic History

The Academic History is mentioned many times for Friends (total 22) and is also common for Acquaintances (total 20), however with less SPT (12 against 15). The SPT mentions decrease considerably (6) respecting Famous, which is supported by significant differences (p<.05) between Famous and Friends. The questionnaire maintains the order of relevance for the different kinds of relationship; yet with very similar values (md=4).

3.4.6 Related People

Between *SPT* and *IND* references *Related People* was a feature that was mentioned almost by every user. *SPT* was more common for *Acquaintances* (12 times against 9 and 10 for *Friends* and *Famous*), which can be explained by the needed to reveal how they were connected to the person (through whom). The questionnaire suggests they find it relevant for *Friends* and *Acquaintances* (md=4) and indifferent for *Famous* (md=3). In fact, there is a statistical significant difference with medium effect between *Friends* and *Famous* (p<.05, r=.40).

3.4.7 Family Information

It was common to talk about the person's family (Figure 3). The little advantage for *Acquaintances* (12 *SPT* against 10 and 8 for *Friends* and *Famous*, respectively) was due to users' justifications on how they knew that person (e.g. "She's Mia's sister"). The questionnaire suggests that participants only considered it relevant for *Friends* (md=4). There were significant differences between *Friend* and *Famous*, with medium effect size (p<.01, r=.45).

3.4.8 Highlights

Highlights were mostly mentioned when talking about Famous (SPT=9; IND=10), which may be the reason why they are famous in the first place. The differences between Famous and Friends are significant (p<.05). This distribution is supported by the questionnaire, where even though it is always relevant, there are significant differences (p<.01) with medium effect size between Famous-Friend (r=.40) and Famous-Acquaintances (r=.47).

3.4.9 City

Participants talked more about the *City* where their friends live (*SPT*=8, *IND*=9), as statistical significance was found between *Friend* and *Famous* (p<.05). In the questionnaire, it was found *Relevant* (md=4) for all types.

3.4.10 Relevant across Relationship Types

There are some attributes which relevance is independent of

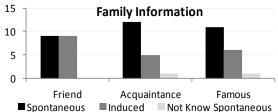


Figure 3. Family Information frequencies per relationship type

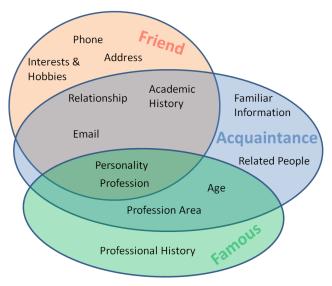


Figure 4. Most relevant attributes per relationship type, based on interviews (*Spontaneous* by more than half users) and questionnaires (median equals 5 – *Very Relevant*)

relationship types. *Profession* and *Personality* are two of the top attributes for every type and *Age*, *Workplace* and *Homepage* are also considered relevant for all of them. *Birth Place* and *Peculiar Expressions* are at the limbo. The former had from 5 to 9 *SPT* references among the different relationship types, suggesting moderate relevance also supported by the questionnaire. The latter had very few *SPT* references (but more IND) but the questionnaire point toward a moderate relevance (md=4).

4. DISCUSSION

Descriptions' length did not vary much either among user demographics or relationship types. Taken at face value, it seems that the amount of information conveyed about people of different types is similar. Still, looking at both the descriptions and the relevance questionnaire we were able to identify the most relevant attributes and the differences among relationship types.

Figure 4 shows the most popular attributes for each relationship type based on the SPT mentions on the descriptions tasks (spontaneously mentioned by more than half participants) and/or the relevance questionnaire (median equals 5 - very relevant). Profession and Personality stand out for being the crossrelationship attributes, evidence supported by both interviews and questionnaires. Other conspicuous characteristics are those that fit only in one type. Interests & Hobbies are prominent when describing Friends and this information is confirmed by the questionnaire (Very Relevant) which shows significant differences between the relationships types; although moderately, it is also relevant for Acquaintances and Famous. The communication characteristics *Email. Phone* and the person's *Address* were rarely resorted to in the descriptions; however they were found Very Relevant. All of them were well rated by the participants for Friends, with an advantage to Phone, which was also the most mentioned (induced) when asked about how they would communicate. For Acquaintances the Email leads and Phone/Address importance decreases. One participant mentioned that he "wouldn't mind to have his phone number, but would contact him by email", which is also suggested by the questionnaire. These characteristics are not relevant for *Famous* people since participants found them unreachable.

Family Information and Related People are prominent to Acquaintances, a fact that arises from the need to explain how they are related to that person. In fact, participants did not realize that need when answering the questionnaire. The dominance of Professional-Related Attributes on the Famous type is also prominent, opposing to the relevance given to Academic History for Friends and Acquaintances. It suggests that when seeking for information about Famous folks, people are more interested in what they are doing now and what they did related to their profession instead of the academic information. The relevance of Academic History for Friends and Acquaintances obtained by the number of SPT references may be slightly influenced by the number of students described in our study.

As for other attributes, *Workplace* and *Homepage* were found relevant for all relationship types, while *Birth Place* and *Peculiar Expression* suggest just a minor relevance. *Highlights*, *Nickname* and *City* tend to be more important for a specific niche. The first one is more relevant for *Famous*, while the others are for *Friends*, as is *Marital Status* with moderate relevance. *Physical Description* revealed to be indifferent and the most unpopular attributes were *Religion*, *Political Ideology* and *Past Addresses*.

5. CONCLUSIONS

We conducted a study that showed that people's personal attributes are not equally relevant in person search contexts. We have scrutinized the attributes relevance when seeking information about friends, acquaintances and famous people and described the differences among these types. In a society where seeking for information about people is a common task, designers can employ these findings to provide a better user experience either by condensing profiles, re-ordering the attributes based on the relationship type, highlighting or even warning about the profile updates related with the most important attributes. To cite one example regarding the Profession attribute, SNS could give more importance when some contact changes his profession, since it is such an important attribute for all relationship types. It is not about what we show to others, but what we want to know about them. The fact that someone shared something does not mean we want to see it and broader results may be burdensome instead of helpful.

We believe we provided valorous insights of attributes relevance in the person search domain; nevertheless we need a higher number of participants in order to argue broad generalization. Moreover, we do not claim that these attributes are the most important features for all contexts. Nevertheless, we believe that in most contexts, the relevant features will be narrowed among the ones this study identified. In particular, probably there are differences in the relevance of attributes between searches about people and alerts about the changes in their profiles. For instance, one may find less interesting to get a friend's workplace when searching for him than to be warned when that attribute changes. It will be one of our major and following concerns, as we will try to validate it and establish the relation between different contexts and relationship types.

6. ACKNOWLEDGMENTS

This work was supported by the Portuguese Foundation for Science and Technology (FCT): individual grant SFRH/BD/66550/2009; project PAELife AAL/0014/2009; and project PEst-OE/EEI/LA0021/2011.

7. REFERENCES

- [1] Fogel, J. and Nehmad, E. Internet social network communities: Risk taking, trust, and privacy concerns, *Computers in Human Behavior* 25- 1, (2009) 153-160.
- [2] Guerreiro, J. and Gonçalves, D. Visualiz'em: "Show me more about him". In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (AVI '12), ACM, New York, NY, USA (2012), 445-448.
- [3] Guerreiro, J. and Gonçalves, D. "Ten things I want to know about you": Unveiling Friends' Attribute Relevance. *In Proc.* of *European Conference on Cognitive Ergonomics* (ECCE'12), (2012).
- [4] Joinson, A. N. Looking at', "Looking up" or "Keeping up with" People? Motives and Uses of Facebook. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'08). ACM, New York, NY, USA (2008), 1027-1036.
- [5] Kelley, P.G., Brewer, R., Mayer, Y., Cranor, L., & Sadeh, N. An Investigation into Facebook Friend Grouping. In Proceeding of Human-Computer Interaction—INTERACT (2011), 216-233
- [6] Lampe, C., Ellison, N. and Steinfield, C., A familiar Face(book): profile elements as signals in an online social network. In Proceedings of the SIGCHI Conference on

- Human Factors in Computing Systems (CHI'07). ACM, New York, NY, USA (2007), 435-444.
- [7] Ozenc, F.K. and Farnham, S.D. Life "modes" in social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11). ACM, New York, NY, USA (2011), 561-570.
- [8] Spink, A., Jansen, B.J., Pedersen, J. Searching for people on web search engines. *Journal of Documentation* 60 (2004).
- [9] Stecher, K., & Counts, S. Thin Slices of Online Profile Attributes. In Proceedings of the Second International Conference on Weblogs and Social Media (ICWSM'08). AAAI Press, (2008), 127-135.
- [10] Wagner, D., Lopez, M., Doria, A., Pavlyshak, I., Kostakos, V., Oakley, I., & Spiliotopoulos, T. Hide and seek: location sharing practices with social media. In *Proceedings of the* 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10). ACM, New York, NY, USA (2010), 55-58.
- [11] Zhou, M., Zhang, W., Smith, B., Varga, E., Farias, M. and Badenes H.. Finding someone in my social directory whom I do not fully remember or barely know. In Proceedings of the 2012 ACM international conference on Intelligent User Interfaces (IUI '12). ACM, New York, NY, USA (2012), 203-206.