DOI:
10.1145/3208039

Document Version
Peer reviewed version

Multiparty Privacy in Social Media

Jose M. Such
Department of Informatics
Faculty of Natural and Mathematical Sciences
King's College London
jose.such@kcl.ac.uk

Natalia Criado
Department of Informatics
Faculty of Natural and Mathematical Sciences
King's College London
natalia.criado@kcl.ac.uk

ABSTRACT

Privacy is not just about what an individual user discloses about herself, it also involves what *her friends* may disclose about her. Multiparty privacy is concerned with information pertaining to several individuals and the conflicts that arise when the privacy preferences of these individuals differ. Social media has significantly exacerbated multiparty privacy conflicts because many items shared are co-owned among multiple individuals. In this paper, we discuss the limited support for multiparty privacy offered by social media sites, the coping strategies users resort to in absence of more advanced support, and current research on multiparty privacy management and its limitations. We then outline a set of requirements to design multiparty privacy management tools.

INTRODUCTION

Over 2 billion users utilise social media to build and participate in Online Social Networks (OSNs), uploading and sharing hundreds of billions of data items [15]. OSNs are, therefore, huge in scale and they are only predicted to keep growing in the forthcoming years both in the number of users and in the amount of data users upload and share. The huge amount of data in social media is user-generated and personal most of the time, which clearly calls for appropriate privacy preservation mechanisms that allow users to benefit from social media while adequately protecting their personal information. Protecting users' privacy is not only essential to respect the Universal Declaration of Human Rights but also to serve as a first-line defence to mitigate cybercrime and other illegal activities that leverage the data obtained due to privacy breaches in social media, such as social phishing, identity theft, cyberstalking, cyberbullying, etc.

There have been many efforts devoted to study privacy in social media and how to protect users' personal information since the very early days of social media such as [10]. However, most of these efforts have focused on privacy from an individual point of view. For instance, advances include research [9] and industry [16] efforts on helping individual users better target their audience by modelling different relationships and social circles beyond the binary friendship model that is prevalent in most social media. While this has indeed helped to advance the state of the art on the topic, the problem of content affecting the privacy of more than one user at the same time has received little attention.

Privacy is not just about what you say or disclose about yourself. It is also about what others say or disclose about you. Evidence shows there are privacy boundaries that are collectively held and managed by individuals within relationships, families, groups, and organisations [23]. With the massive growth of Social Media, however, collectively held privacy boundaries have become extremely challenging to maintain, as many of the hundreds of billions of items uploaded to social media are co-owned by multiple users [15][14], yet mainstream social media only allow the user uploading a co-owned data item to set its privacy settings, which often leads to conflicts and severe privacy violations [35][33]. Multiparty privacy (MP) aims to facilitate the coordination of collectively held privacy boundaries by all individuals that co-own a data item online, as the privacy of all of them may be at stake depending on with whom the co-owned data item is shared[1]. MP particularly focuses on supporting the detection and resolution of multiparty privacy conflicts (MPCs), when individuals whose privacy may be affected by the same co-owned data item have conflicting privacy preferences. Take a simplified but illustrative example of MPC: Alice takes a photo of her and Bob. Mainstream social media would only allow Alice (assuming she uploads the photo) to set the privacy settings for the photo, but what if Bob would not like to share it with some of the friends Alice would like to share the photo with? MP is concerned with not only photos but also other social media content such as posts, videos, comments, events, etc. Beyond social media, MP could also be useful in other social computing domains, in which information is co-created and co-owned by multiple users, so all these users should have a say on to whom this information is shared, such as collaborative software (e.g. cloud-based collaborative documents), cloud-based file sharing [24], internal/external wiki pages, blogs, collective intelligence, crowdsourcing, etc.

Designing MP tools is a complex and difficult task, as users have different privacy attitudes and preferences; they socialise online with multiple types of relationships; and they share varying amounts of different types of content. In this paper, we discuss the limited MP support users have nowadays, the coping strategies users are forced to resort to in the absence of adequate MP support, and the latest developments in MP mechanisms and tools. Based on this, we outline a roadmap for future research with a set of requirements for developing MP tools.

SOCIAL MEDIA SUPPORT FOR MP

Mainstream social media sites support some sort for MP, which mainly comes in the form of two mechanisms: (i) tagging/untagging; and (ii) reporting inappropriate content.

Tags are normally used to name people that appear in a photo with a link to their profile. People tagged in a photo can, however, untag themselves from the photo. There are some social media sites (like Facebook) in which you can opt-in to receive

[1] MP is different from other collective approaches that focus on protecting just *one* individual, e.g., [4],[6],[21].

---
* This is the authors' version of the paper accepted for publication in CACM, 2017.

notifications about the photos you have been tagged in to approve tags before they become effective. Tagging/untagging represents some sort of MP, but it has three main limitations. The first limitation is that even if you untag yourself from a photo before anyone seeing it, this does not mean that your friends will not end up seeing the photo anyway. For instance, Alice and Bob are in a photo that Alice uploads to Facebook tagging Bob in it. Bob receives a notification, he revises the photo and decides not to approve the tag, e.g., because he feels embarrassed about the photo. The point is that the photo, even without Bob being explicitly tagged, will be shared according to what Alice decides. That is, if Alice decides to share with her friends, and Alice and Bob share some friends, all these friends will be able to see the photo in Alice's wall anyway. The second limitation is that tagging/untagging is supported for photos but not for other items such as posts, comments, events. Posts and comments do usually have the option to include *mentions* (using special symbols such as '@'), but these mentions are only controllable by the post/comment creator – though users can remove comments to their posts/photos. Finally, many users state that they feel very uncomfortable untagging themselves from photos because it may offend (from a social angle) the person who tagged them in the photo [2].

Regarding reporting, most social media sites allow users to report when content published by others is not appropriate. This mechanism is mainly used to deal with *highly* inappropriate (or even illegal) content such as nudity, hate speech, violence, and other very serious offences. After being reported, the provider decides unilaterally what to do with the content (delete it or not). Although this mechanism is of utmost importance to fight against these very serious offences, it is not appropriate for all MP scenarios, as there are many cases in which privacy violations can happen without necessarily being related to these offences. For instance, it may just be the case that you are not comfortable sharing some information with some other people, or you want to conceal information from your work colleagues, etc. Also, it is important to highlight that reporting is only a *reactive* mechanism, which only activates after content has already been published and someone flags it as inappropriate. However, when the content is flagged, it may well be too late, the privacy violation may have already happened and the derived consequences may be unrecoverable, or other users may have been able to download the content and distribute it using other channels.

The problem of MP starts being recognised by mainstream social media as demonstrated by the latest re-vamp of Facebook's privacy controls[2]. In particular, Facebook's Privacy Basics now explains the newly-introduced option to contact users about photos you do not like. The mechanism works as follows, if a user is tagged in a photo and she does not like the photo, she can now flag the photo as not liking it, which then opens up a message window containing a form with the recipient field set to the one who uploaded the photo, so that the user who does not like the photo can ask the user who uploaded it to remove it and include an optional reason for the removal. Although this is a step forward which very much recognises the issue of MP, it still falls short because of multiple reasons, some of them also shared with tagging/untagging and reporting inappropriate content: (i) the process happens once the photo has already been published, so any potential privacy breaches may have already occurred; (ii) it takes time to take down a photo that has already been published -- e.g., Lian et al. [19] calculated the time it takes for a photo URL to become unavailable after having deleted the photo from the social media site, which turned out to be 3 days on Instagram, 7 days on Facebook, 14 days on Flickr and over 30 days in MySpace and Tumblr; (iii) it does not enable *collective* negotiation, as the photo may involve other people and not only the one who uploaded it and the one who complains about it; (iv) everything needs to be done manually, which introduces an unbearable burden on the users considering the large amount of friends users have online; and (v) this mechanism has only been implemented for photos but not for other types of content such as posts, comments, events, etc.

[2] https://www.facebook.com/about/basics/how-others-interact-with-you/

USER COPING STRATEGIES FOR MP

As shown above, there is a distinct lack of built-in capabilities in current social media infrastructures to help users compromise by actively negotiating with others [40]. Users are forced to communicate outside social media and apply a number of *coping strategies* to try to overcome or work around that lack of technical support. Basically, most of these coping strategies consist of actions or behaviours in the offline world that aim to prevent MPCs from happening online. Research uncovered several examples of these coping strategies, which very much stress the need for MP tools. We discuss some examples of coping strategies and their shortcomings next (summarised in Table 1).

Table 1. Examples of coping strategies

Strategy	***Main Drawbacks***
Try to anticipate consequences for others [18]	Impossible to always anticipate privacy consequences.
Seek approval before posting [18]	Too much burden on the user that uploads the item.
Inside jokes and cloaking [3]	It does not scale and it is not feasible for some types of content.
Alternative sharing media [2]	MPCs can happen in other media too. Also, one user cannot control which media others use to share content.
Change offline behaviour when cameras around [2][18]	Very difficult due to the pervasiveness of smart phones and wearables.
Negotiation of a shared policy with other users [40][18][2]	It could easily become a burden on the users due to the amount of co-owned content.

One of the offline strategies people utilise before posting an item to avoid MPCs is trying to anticipate whether the item could be sensitive to anyone potentially affected by it [18]. For instance, if Alice and Bob appear together in a photo but Bob appears clearly inebriated, then it is likely that Alice may consider this by either not posting the photo or sharing it only with a restricted number of friends. However, this does not always work, as sometimes the person to post an item cannot anticipate the consequences this may have for others beforehand. An example is given in [18], a person was congratulated by a friend about being accepted for a Masters programme via a comment, but the person had to quickly remove the comment as he had not yet told his employer about it and his employer was also friend of his online. Note even if the

person removed the comment quickly, there was still the risk that his employer may have already noticed the comment before being removed.

Users sometimes ask the other co-owners of an item for approval before sharing it [18]. The problem with this strategy is that it is done offline without any technical means that could facilitate this. That is, one would need to ask permission offline to all people that may be affected by each and every item they upload. Also, when somebody was not ok, they would need to negotiate a solution (e.g., reduce the initial audience, decide not to upload, etc.). This would quickly become an unbearable burden on users.

It has also been observed that teens cloak their messages and share photos with inside jokes [3]. For instance, Boyd and Marwick [3] report an example of a girl writing a post in Facebook about something she knew only her close friends would understand, as she wanted to prevent other friends from knowing what she actually meant. The downside of this strategy is that it clearly does not scale and may not be feasible for all photos or other types of items that people would like to share. For example, a photo about your travel to Mauritius cannot be easily cloaked in case you want to share it with some people but not with others.

As social media proves inadequate to manage disclosures in MP scenarios, some users switch media to share content using other technologies such as cloud-based file sharing, instant messaging, or e-mail attachments [2]. This has the advantage of protecting not only their own content but also limiting the privacy risks for others. There are, however, three main disadvantages as well. Firstly, this may be possible for photos, videos, etc. but not for other types of content such as events, comments, etc. Secondly, users cannot control which technologies their friends use; i.e., their friends could still upload photos using social media without users being able to do anything about it. Thirdly, these technologies might also lead themselves to MPCs. For instance, one user may share a video in a Whatsapp group in which there are people with whom other users in the video would not like to share it.

Users also confirmed that, in the absence of better ways to manage MP situations, they actually change and tightly control their offline behaviour. For example, people behave in a different way when they see a camera around [2][18]. If you know a friend likes to take photos and posts them very often, you may decide not to hang out with her to avoid any undesired photos being posted. This highlights the extent to which people feel unable to participate in MP decisions. The effectiveness of this strategy is again very limited, mainly due to the pervasiveness of smart phones and wearable devices, being always alert and constantly modifying your offline behaviour is infeasible.

One of the most interesting strategies perhaps is that users collectively negotiate and achieve offline agreements and compromises about what gets posted and to whom it gets shared [40][18][2]. For instance, a group of friends could agree that the photos they take in a trip can only be shared among them or with close friends of them. Interestingly, it turns out users are always very open to consider and accommodate others' preferences as much as possible [40][18]. In addition, research uncovered that users do not want to cause any deliberate harm to their friends and will normally listen to reasonable objections, which also acts as a way of reaffirming and reciprocating relationships [40]. The main problem with this strategy, as with many of the other strategies seen so far, is that it does not scale. It is impossible for users to be constantly negotiating with hundreds of friends about hundreds of photos without technical aid.

RESEARCH ON MP TOOLS

It seems clear considering all the cases above that users actively seek to work around the problem of not having adequate technical support for MP. However, the effectiveness of the coping strategies they use for this seems rather limited according to the drawbacks these strategies have. This has inspired researchers to design interfaces and computational methods that empower users to collectively manage MP in more effective and efficient ways than the current coping strategies they are nowadays forced to resort to. Although research on this area is still in its infancy, there have been a number of proposals that we categorise below into 5 main approaches (summarised in Table 2), highlighting their strengths and limitations. Note that other works in addition to those discussed have also been published but we could not include all of them due to the space and maximum references allowed, and have instead included those we considered the most representative of each approach.

Manual approaches. The first research stream proposed support for MP by helping users to identify where MPCs can or did occur [2][39]. For instance, [39] presents a way to specify strong and weak sharing preferences so that these preferences could be inspected to find conflicts. Also, [2] introduces a system whereby users tagged in a photo can contact the user who uploaded the photo to ask to remove it or to restrict the audience of the photo, which resembles the functionality Facebook introduced some time later [7]. While these approaches represented a stepping stone, recognised the problem of MP, and proposed a partial solution to it; they left all the negotiation process to resolve detected conflicts to happen without any particular technical aid. That is, users need to resolve every potential MPC in a *manual* way, which may become an unbearable burden considering the massive amount of content uploaded and the number of friends that users have in social media.

Auction-based approaches. Another research stream proposed solving potential MPCs using a bidding mechanism [30]. Users bid for the sharing decision they would prefer the most and the winning bid determines the sharing decision that will be taken for a particular item. These approaches were the first ones to consider a semi-automated method to aid users in collectively defining a sharing decision – e.g., the outcome of the auction is computed automatically from the bids users specify. However, users may have difficulties to comprehend the mechanism and specify appropriate bid values in auctions, and users are required to bid for each and every item co-owned with others.

Aggregation-based approaches. These approaches suggest a solution to a MPC by aggregating the individual privacy preferences of all users involved. They can be abstractly conceptualised as voting mechanisms, where the preferences of each user affected by an item count as one vote (sometimes weighted) for sharing/not sharing. Then, a voting rule models how each of these mechanisms aggregates votes together. For instance, in majority voting [5], the preference of the majority of users is taken as the decision to be applied to the content. Another example would be veto voting [35], so that if there is one of the users affected by the content who opposes sharing, then the content is not shared. The main problem with these approaches is that they always aggregate preferences in the very same way. For instance, using majority voting always means that even when content can be very sensitive and lead to privacy violations for one user, it will be shared if the majority of users wishes to. In contrast, always using veto voting may be too restrictive and impact the known benefits users get from sharing in social media

[29]. Subsequent works [12] do recognise this issue and consider more than one way of aggregating user preferences. However, it is up to the one who uploads the item to decide the aggregation method to apply. This requires the user who uploads the item to anticipate the consequences for others, which may be a very difficult task as discussed above, and it may not always render the optimal solution.

Table 2. Summary of MP approaches with example references

Approach	*Short Description*	*Main Drawbacks*
Manual [2][39]	Users are provided with a way of detecting MPCs, and they can manually resolve them when detected.	It may easily become a burden on the users, as they do not provide automated support for conflict resolution.
Auction-based [30]	Users gain fictitious money that they can invest in auctions bidding for the most desired sharing decision for co-owned items.	Users may have difficulties to understand and manage the process appropriately.
Aggregation-based [5][35][12]	Individual privacy preferences of all users are aggregated using a rule or set of rules to produce a joint sharing decision.	Individual privacy preferences are aggregated in the same way or the uploader chooses the method to aggregate.
Adaptive [32]	Different situations are modelled based on a number of factors and a different sharing decision is suggested depending on the situation.	It is difficult to model all possible factors that determine a situation and the best method to achieve an optimal sharing decision.
Game-theoretic [13][31][17][25]	Users or automated software agents negotiate a solution following an established protocol. Both the protocol and the negotiation strategies are analysed using game-theoretic solution concepts.	Users' behaviour in social media seems not to be perfectly rational as there are many very social idiosyncrasies that play a role in MP.
Fine-grained [14][36]	Users define individualised access control decisions over personally-identifying objects within a photo, e.g. users deciding whether or not their faces are blurred.	Blurring objects (e.g. faces) within a photo may not be the optimal solution in terms of the utility of the information shared and/or protecting users' privacy.

Adaptive approaches. These approaches automatically *infer* the best way to solve a MPC based on the particular situation [32]. These approaches model a situation considering factors such as the individual preferences of each user, the sensitivity of the content, the relationships to the potential audience, etc. Then, a particular situation instantiates particular *concessions* that are known to happen when people negotiate offline an agreement about sharing co-owned items [40][18][2]. Thus, these approaches automatically *adapt* to the situation at hand, turning as restrictive as veto voting if the situation requires so (e.g., if the item is very sensitive), or suggesting sharing in other situations (e.g., someone having special interest in sharing and the others not caring much about it). While these approaches capture the known situations of when concessions happen during offline negotiations, it is difficult to model all possible situations, and they may not capture opportunistic concessions or agreements that may arise in potentially unknown situations.

Game-theoretic approaches. Another approach has been to define negotiation protocols, which are a means of standardising the communication between participants in the process of negotiating a solution to a MPC by defining how the participants can interact with each other. These protocols are then enacted by users themselves manually [15] or automatically by software agents [31][17] to negotiate an agreed sharing decision for a particular item. Participants can follow different strategies when enacting the negotiation protocols, and these strategies are analysed using well-known game-theoretic solution concepts such as the Nash equilibrium. This allows, for instance, to determine analytically which are the best strategies that participants can play as well as to find strategies that are stable (strategies in which no participant has anything to gain by changing only her own strategy unilaterally). While these proposals provided elegant frameworks from a formal point of view and build upon well-studied analytic tools, they may not work well when used in practice [13]. This is because users' behaviour does not seem perfectly rational in practice (as assumed in these approaches), and even if some are starting to consider other factors like reciprocity [17] and social pressure [25], they are still far from considering the many very social idiosyncrasies that play a role in MP [18][40].

Fine-grained approaches. The last research stream focuses on preventing MPCs by allowing each user in a photo to independently decide whether some personally-identifying objects within the photo are shown or blurred [14][36]. In particular, one of the first works in this approach allowed users to individually decide whether their face is shown or blurred [14]. The process works as follows: 1) the users in a photo are identified using face recognition algorithms such as Facebook's DeepFace algorithm [34]; 2) the users recognised are notified and they can suggest the list of friends who can have access to the photo; 3) when a user wants to access a photo, she will only see the faces of the users that have granted access to her and the other faces in the photo will appear blurred. However, blurring faces (or other objects in a photo) may impact the utility of the photo being shared, negatively impacting the benefits people get by sharing in social media [29], and there is also the risk that a person can be re-

identified even if her face (or other objects in a photo) has been blurred [22]. Hence, when a collaboratively-agreed solution to a MPC is possible, that solution might be more desirable than enforcing access separately, as the photo will not lose any utility (no object blurred) but the audience of the photo will be negotiated to remove access to any undesired people.

REQUIREMENTS FOR MP TOOLS

Building upon the previous analysis on existing approaches and their limitations, we now outline a set of requirements to develop MP tools that empower users to collectively manage their privacy together with others and overcome these limitations. These tools would aid end-users to identify potential MPCs and, when MPCs are identified, provide support for their resolution (e.g. in the form of recommendations), allowing an appropriate "*boundary regulation process by actively negotiating one's boundaries with others*" [40]. Next, we describe with more detail each of the requirements.

Design informed by real-world empirical data

None of the existing approaches are grounded on a deep understanding of MPCs and their optimal solution in practice. This is in part due to not having enough empirical evidence about MPCs yet. Such an empirical base is utterly essential to inform the design of MP tools that overcome the limitations identified in the existing literature. As mentioned above, researchers have shed light on how users are forced *online* to resort to coping strategies to work around the lack of appropriate support for MP [3][18][40][2], and there is evidence of how collectively held privacy boundaries are managed *offline* [23]. While this previous research already provides a very good foundation to build upon, further research is needed to better understand when and how often MPCs actually happen online and, more importantly, when they become a problem or lead to potential privacy violations and hence need a solution. Particular instances of MPCs users faced could be studied to understand whether they happened despite coping strategies being used, how users came up or would come up with the optimal solution for the MPCs studied, and the factors that played a role in the process. Some very recent research goes in this direction [33], having contributed the first empirical and public dataset of MPCs. Having this empirical base about MPCs would ultimately underpin a thorough understanding of MPCs and the nuanced factors that affect them from the ground up, which could then be used as the basis to design MP tools that offer support to different types of users, social groups and relationships and can recommend optimal solutions to MPCs. Recent efforts on privacy engineering should be leveraged to easy the challenging task of going from empirical evidence to privacy design [11].

User-centric MP controls

The main challenge here is how to develop usable MP tools in line with the empirical base mentioned above, so users could effectively manage MP with minimal effort. However, MP tools should aim for usability without becoming a *fully* automated solution, as this may not achieve satisfactory results when it comes to privacy in social media. Instead, users may have to provide some input into MP tools, which will then provide a *recommendation*, as very recent research has shown that the optimal solution for an MP conflict could be predicted given some input from the users, like the reason for their preferred privacy policy [8]. However, if users have to intervene to express their individual privacy preferences and/or to accept/decline the solution recommended for each and every co-owned item and potential conflict, would this not easily become a burden on the users? How do we find adequate trade-offs between intervention and automation? There are previous studies on individual privacy in social media that could help for this: i) tools like AudienceView [20] could be used to show and/or modify the suggested solution or express individual preferences; ii) approaches similar to [7] could be used to *learn* the way users respond to MP over time; and iii) approaches like [38] could be used to create suitable defaults for MP settings.

Scaled-up and comparable evaluations

The existing approaches for MP presented above were either not evaluated empirically with users ([5][17][25][30][31][39]) or the user studies conducted were low-scale with at most 50 participants ([2][12][13][14][32][36]). This is in part due to a distinct lack of systematic and repeatable methods and/or protocols to evaluate MP tools and compare them to each other. In order for evaluations to be more conclusive and generalizable, MP tools should be evaluated considering wider and more varied populations. Also, evaluation protocols should be developed with a view to maximise *ecological validity,* which is particularly challenging in this domain. Firstly, participants in user studies would always seem reluctant to share sensitive information with researchers [37] (e.g., photos they feel embarrassed about and prefer not sharing online), which would bias any evaluations towards non-sensitive issues only, leaving out the scenarios where the adequate performance of MP tools would be critical. An alternative could be evaluations with fake data/scenarios where participants self-report how they would behave, but the results may not match participants' actual behaviour in practice due to the well-known dichotomy between privacy attitudes and behaviour [1]. Secondly, conducting MP evaluations *in the wild* is very difficult, as it would require all the users affected by a particular piece of content to be studied together to understand the conflicts and whether the solutions to the conflicts are optimal. A possible way forward could be methodologies based on *living labs*, which would integrate and validate research in evolving real life contexts.

Privacy-enhanced party recognition

Given a particular item uploaded, MP tools should derive the users who are affected by the item. For instance, if a user uploads a photo and tags in it all the other users that appear in the photo, MP tools can directly use this to know which users are involved. However, users many times either do not tag all people clearly identifiable in a photo or incorrectly tag people who actually do not appear in the photo. Face recognition software could be used for this, such as the one developed by Facebook researchers called DeepFace [34], which has 97.35% accuracy. The question that arises is whether using face recognition software could be too privacy invasive for individuals, i.e., the social media provider would be able to identify individuals in any photo even for photos outside the social media infrastructure, or individuals could be misidentified and wrongly associated with items that are not relevant to them (note even if accuracy of face recognition is high and false positives are low, the number of items and users is huge). Interestingly, this seems to open a completely new and exciting type of privacy-related trade-off compared to the well-known privacy-utility trade-off, which would be multiparty vs. individual privacy. Note, however, that a multiparty-individual privacy trade-off will not be needed if privacy-preserving face recognition methods [27] are used by MP tools, so that parties would be recognised while preserving their privacy. Beyond

photos, party recognition may be easier for some content type such as events (people invited or attending are explicitly mentioned) or even more challenging for some other content such as text posts, in which affected users may not always be explicitly tagged.

Support for inferential privacy

Another issue not considered before in a MP context is that of inferential privacy. That is, it may not only be about what your friends say about you online, but also what it may be inferred from what they said regardless of the type of content. For instance, Sarigol et al. [27] have demonstrated the feasibility of constructing shadow profiles of sexual orientation for users and non-users, using data from more than 3 Million accounts of a single OSN. Note that negotiations or agreements for the case of inferential privacy may be more complex, as the reasons not to publish content may not be about the content itself but more about the consequences in terms of the information that may be inferred from it, so solutions to this type of MPC might be more difficult to comprehend by users, which would also challenge the usability and understandability of MP tools. Also, we are unaware of any social media site that provides users with any sort of controls for inferential privacy; let alone any research conducted that considers both MP and inferential privacy together.

Privacy-preservation guarantees

Last but not least, MP tools should provide some sort of individual privacy guarantees. This is particularly important when a multiparty agreement is not possible. For instance, a user may be posting on purpose content that defames another user. In these cases, there may be room for enforcing individual privacy preferences to some extent. For instance, a possible solution for photos is the work already mentioned above [14], which would allow users to control whether their face is shown or blurred in a particular photo. This seems an appropriate solution when a MP conflict arises and no agreement is found by the users affected, so instead of the *winner taking it all*, the outcome is that all users affected are guaranteed their individual privacy to some extent. This, however, does not completely remove the identification risks, as acknowledged by the authors of [14], because there is still the chance the user may be recognised even after her face has been blurred [26], and approaches that are able to remove the full body of a person and reconstruct the image are still not there, though there are approaches that already recognise user's body/gesture [28].

CONCLUSION

Multiparty privacy (MP) is an important problem in social media that also expands into other areas of social computing where there is co-owned information such as blogs, collective intelligence, wiki pages, and cloud-based file sharing [24] and collaborative documents, which have received even less attention when compared to social media for this matter. As highlighted in this paper, mainstream social media do not provide adequate support for MP and, as a result, users are forced to use different coping strategies that are far from optimal. Thus, there is a need for the development of novel privacy-enhancing techniques and mechanisms to help users to manage MP. We still have a long way to go to make such mechanisms a reality and embed them in highly usable tools ready to be utilised by end-users, partly due to the complex nature of MP and social behaviours, which requires an interdisciplinary approach to MP. In this paper, we have introduced the area of MP tools, discussed its current state and advances, and defined a set of requirements to shape the agenda for future research in this area.

ACKNOWLEDGMENTS

We would like to thank the EPSRC for supporting this research under grant EP/M027805/1. We would also like to thank William Aiello and the anonymous reviewers for their very useful and helpful comments on a previous version of this manuscript.

REFERENCES

[1] Acquisti, A., & Gross, R. (2006, January). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy enhancing technologies PET* (pp. 36-58).

[2] Besmer, A., & Richter Lipford, H. (2010). Moving beyond untagging: photo privacy in a tagged world. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1563-1572).

[3] Boyd, D., & Marwick, A. (2011). Social steganography: Privacy in networked publics. International Communication Association, Boston, MA.

[4] Calikli, G.; Law, M.; Bandara, A. K.; Russo, A.; Dickens, L.; Price, B. A.; Stuart, A; Levine, M. and Nuseibeh, B. (2016). Privacy Dynamics: Learning Privacy Norms for Social Software. In: 2016 IEEE/ACM 11th International Symposium on Software Engineering for Adaptive and Self-Managing Systems, Association of Computing Machinery pp. 47–56.

[5] Carminati, B., & Ferrari, E. (2011) "Collaborative access control in online social networks," in IEEE CollaborateCom, pp. 231–240.

[6] Criado, N., & Such, J.M. (2015). Implicit contextual integrity in online social networks. *Information Sciences*, 325:48–69.

[7] Fang, Lujun, and Kristen LeFevre. "Privacy wizards for social networking sites." Proceedings of the 19th international conference on World wide web. ACM, 2010.

[8] Fogues, R. L., Murukannaiah, P. K., Such, J. M., & Singh, M. P. (2017). Sharing policies in multiuser privacy scenarios: Incorporating context, preferences, and arguments in decision making. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *24*(1), 5.

[9] Fong, P. W. Relationship-based access control: protection model and policy language. In Proceedings of the first ACM conference on Data and application security and privacy (pp. 191-202), 2011.

[10] Gross, R., & Acquisti, A. (2005). Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM workshop on Privacy in the electronic society (pp. 71-80).

[11] Gürses, S., & del Alamo, J. M. (2016). Privacy engineering: Shaping an emerging field of research and practice. *IEEE Security & Privacy*, *14*(2), 40-46.

[12] Hu, H., Ahn, G. J., & Jorgensen, J. (2013). Multiparty access control for online social networks: model and mechanisms. *IEEE Transactions on Knowledge and Data Engineering 25*(7), 1614-1627.

[13] Hu, H., Ahn, G. J., Zhao, Z., & Yang, D. (2014). Game theoretic analysis of multiparty access control in online social

networks. In *Proceedings of the 19th ACM symposium on Access control models and technologies* (pp. 93-102).

[14] Ilia, P., Polakis, I., Athanasopoulos, E., Maggi, F., & Ioannidis, S. (2015). Face/Off: preventing privacy leakage from photos in social networks. In *Proceedings of the 22nd ACM Conference on Computer and Communications Security (CCS)*.

[15] Internet.org (2014). A focus on efficiency. http://internet.org/efficiencypaper (last visited Jan., 2016).

[16] Kairam, S., Brzozowski, M., Huffaker, D., & Chi, E. (2012). Talking in circles: selective sharing in google+. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1065-1074). ACM.

[17] Keküllüoğlu, D., Kökciyan, N., & Yolum, P. (2016). Strategies for privacy negotiation in online social networks. In *Proceedings of the ACM International Workshop on AI for Privacy and Security (p. 2)*.

[18] Lampinen, A., Lehtinen, V., Lehmuskallio, A., & Tamminen, S. (2011). We're in it together: interpersonal management of disclosure in social network services. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 3217-3226). ACM.

[19] Liang, K., Liu, J. K., Lu, R., & Wong, D. S. (2015). Privacy Concerns for Photo Sharing in Online Social Networks. IEEE Internet Computing, 19(2), 58-63.

[20] H. R. Lipford, A. Besmer, and J. Watson, Understanding privacy settings in facebook with an audience view, in *UPSEC*. USENIX, 2008, pp. 1–8.

[21] Nissenbaum, H. (2004). Privacy as contextual integrity. *Wash. L. Rev.*, *79*, 119.

[22] Polakis, I., Ilia, P., Maggi, F., Lancini, M., Kontaxis, G., Zanero, S., ... & Keromytis, A. D. (2014, November). Faces in the distorting mirror: Revisiting photo-based social authentication. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (pp. 501-512). ACM.

[23] Petronio, S. (2012). Boundaries of privacy: Dialectics of disclosure. Suny Press.

[24] Ramokapane, K. M., Rashid, A., & Such, J. M. (2017). " I feel stupid I can't delete...": a study of users' cloud deletion practices and coping strategies. *USENIX Symposium on Usable Privacy and Security (SOUPS)*, 241-256.

[25] Rajtmajer, S., Squicciarini, A., Such, J. M., Semonsen, J., & Belmonte, A. (2017). An Ultimatum Game Model for the Evolution of Privacy in Jointly Managed Content. In *International Conference on Decision and Game Theory for Security (GameSec)* (pp. 112-130).

[26] Sadeghi, A. R., Schneider, T., & Wehrenberg, I. (2010). Efficient privacy-preserving face recognition. In *Information, Security and Cryptology–ICISC 2009* (pp. 229-244). Springer Berlin Heidelberg.

[27] Sarigol, E., Garcia, D., & Schweitzer, F. (2014, October). Online privacy as a collective phenomenon. In *Proceedings of the second ACM conference on Online social networks* (pp. 95-106).

[28] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., ... & Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, *56*(1), 116-124.

[29] Spiekermann, S., Krasnova, H., Koroleva, K., & Hildebrand, T. (2010). Online social networks: why we disclose. *Journal of Information Technology*, *25*(2), 109-125.

[30] Squicciarini, A. C., Shehab, M., & Paci, F. (2009, April). Collective privacy management in social networks. In *Proceedings of the 18th international conference on World Wide Web* (pp. 521-530). ACM.

[31] Such, J. M., Rovatsos, M. (2016). Privacy Policy Negotiation in social media. ACM Transactions on Autonomous and Adaptive Systems 11(1):1-29.

[32] Such, J. M., Criado, N. (2016). Resolving Multiparty Privacy Conflicts in social media. IEEE Transactions on Knowledge and Data Engineering 28 (7):1851-1863.

[33] Such, J. M., Porter, J., Preibusch, S., & Joinson, A. (2017). Photo Privacy Conflicts in Social Media: A Large-scale Empirical Study. In *Proceedings of the 2017 ACM CHI Conference on Human Factors in Computing Systems* (pp. 3821-3832).

[34] Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014 (pp. 1701-1708).

[35] Thomas, K., Grier, C., & Nicol, D. M. (2010,). Unfriendly: Multiparty privacy risks in social networks. In *Privacy Enhancing Technologies* (PET) (pp. 246-252). Springer Berlin Heidelberg.

[36] Vishwamitra, N., Li, Y., Wang, K., Hu, H., Caine, K., & Ahn, G. J. (2017). Towards pii-based multiparty access control for photo sharing in online social networks. In *Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies* (pp. 155-166).

[37] Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., & Cranor, L. F. (2011). I regretted the minute I pressed share: A qualitative study of regrets on Facebook. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (p. 10).

[38] Watson, Jason, Heather R. Lipford, and Andrew Besmer (2015). "Mapping user preference to privacy default settings." *ACM Transactions on Computer-Human Interaction (TOCHI)* 22.6: 32.

[39] Wishart, R., Corapi, D., Marinovic, S., & Sloman, M. (2010). Collaborative privacy policy authoring in a social networking context. In IEEE international symposium on Policies for distributed systems and networks (POLICY) (pp. 1-8).

[40] Wisniewski, P., Lipford, H., & Wilson, D. (2012). Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 609-618).