BENJAMIN GRESCHBACH AND SONJA BUCHEGGER

Friendly Surveillance – A New Adversary Model for Privacy in Decentralized Online Social Networks

In pace with the ever increasing popularity of Social Network Services (SNS) the critical privacy flaws of these applications got into focus of media as well as research interest in the last decade. The centralized aggregation of personal user data has been identified as a fundamental problem of popular services such as Facebook or Google+.

To mitigate this shortcoming the concept of a Decentralized Online Social Network (DOSN) has evolved, where users form a peer-to-peer (P2P) network to corporately operate the service. While this architectural shift immediately eliminates the threat of a central provider adversary, new challenges to protect the users' privacy arise.

In this paper we focus on the friend adversary model – that is an attacker that exploits the social relationship status established to the target user. We examine the properties of a friend adversary in a decentralized system by analyzing its capabilities, attack impacts as well as incentives and compare the results to the centralized case. We identify several implementation issues of DOSNs that can alleviate illegitimate data collection for a friend adversary. Furthermore, background knowledge about a user may complement this information to mount relevant and privacy invading attacks. We conclude that friend adversaries can be powerful attackers indeed and propose to consider this hitherto less emphasized threat for DOSN implementations.

I. Introduction

The number of users of social networking services (SNS) online has grown fast over the last years, currently the estimate for Facebook, to take the most popular example, is exceeding 800 million users. In parallel with the increase of users and amount of data stored about these users, there have been growing concerns about user privacy in SNS. Prompted by such concerns, there is now substantial research to investigate the extent of privacy threats and to propose more privacy-friendly alternatives. In some cases, public pressure lead to changes of terms-of-service of existing SNS. Privacy issues in SNS are well known and reported in both research papers and news media. The main threat to user privacy stems from the massive collection of private, sensitive and personally identifiable information. These data can be mined for targeted advertisements by the SNS provider or, more importantly, leak to third parties that the user has no agreement with. These leaks can happen either intentionally or by accident. In either case they represent a loss of control, users cannot decide or even know who can see their data. Several prominent leakages have been reported, not only from social networks, but also other centralized data repositories for health data, credit-card or user account information. To address the dual privacy threat of an SNS provider, namely data mining and leakage risk, it has been proposed to decentralize the control of SNS. Decentralized online social networks (DOSN) have become a lively branch of research.

While decentralization rids the system of a privacy-threatening bottleneck, it also distributes some of the power of a (logically) centralized provider to other players in the system and there is no notion of a single entry point or reference monitor in the system anymore. New types of adversaries become more relevant as the threat of the provider decreases. On the system level, sniffers can analyze traffic better than before, as the destination itself can now lead to inferences about data and social relations. On the social network level, friends can combine new inferences from more accessible metadata with their personal background knowledge and thus become potentially powerful adversaries for targeted attacks. In this scenario, the threat model shifts to hitherto less expected privacy leakages.

In this paper, we focus on the adversary type of friends in decentralized online social networks, as they become the most powerful potential attackers after the removal of the SNS provider. Although the threat from the latter is mitigated by decentralization, it is only a first step toward privacy-preserving social networks and decentralized services need to be designed with appropriate adversary models in mind to avoid privacy breaches. We analyze the potential of friend adversaries in a DOSN setting and find that friends can be powerful attackers indeed.

The paper is organized in the following way: After referring to related work in Section II., we discuss the concept of Decentralized Online Social Networks in Section III. Next, the friend adversary model and its properties are analyzed in Section IV. Section V. concludes with comparing the adversary model in different social network architectures, discussing the new challenges arising from it and sketching possible countermeasures to mitigate these novel threats.

II. Related Work

The impact of SNS on their users' privacy has been extensively studied. Gross et al. (2005) have identified several threats of SNS usage such as stalking; de-anonymization of external sensitive sources (e. g. anonymized medical records); identity theft (e. g. by social insurance number reconstruction); user profiling by building a digital dossier and simplified social engineering. Paul et al. (2011) underline the consequences of a massive central data aggregation in conjunction with an advertising-based business model of major SNS providers. They warn against the possibilities of direct misuse or unintended leakage of this data that is not appropriately protected and hard to anonymize. Krishnamurthy and Wills (2010) show that relevant leaks of personal information do occur in practice. Besides the suggestion to use more suitable data security such as state of the art cryptography, one main approach to address the privacy issues is the decentralization of the SNS. Buchegger et al. (2009) propose the

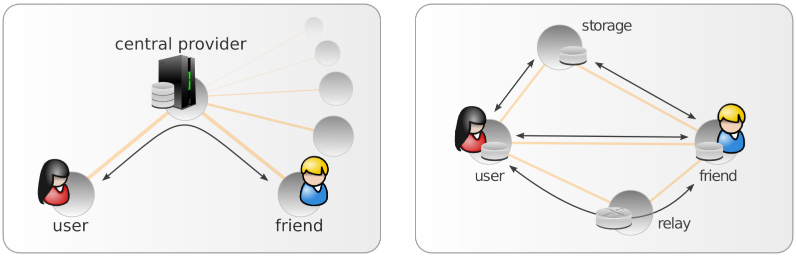
PeerSoN system where (encrypted) content data is distributed using a P2P network formed by the users of the SNS. Thus, the single point of failure of a centralized system is removed and the users' ownership of their data preserved. Aiello and Ruffo (2010) elaborate on a Distributed Hash Table (DHT) based architectural framework supporting SNS functionality. They propose authentication on the routing level to defend against common attacks against the overlay and discuss implementations of SNS requirements such as access control, reputation management and search operations. Cutillo et al. (2009) introduce *Safebook*, an architectural approach focusing on communication anonymization. Content is stored at trusted friend nodes and requests are routed through a mix-network formed by social links to obfuscate information flow. The *Persona* project (Baden et al., 2009) proposes the use of an attribute-based encryption scheme to realize group encryption without encrypting the symmetric content key with the public keys of all recipients. Finally, Bodriagov and Buchegger (2011) scrutinize several proposals for DOSN tailored encryption schemes and evaluate their performance for different SNS operations.

III. Decentralized Social Network Services

Decentralized Online Social Networks (DOSN) are evolving as a promising approach to mitigate design-inherent privacy flaws of logically centralized services like Facebook, Google+ or Twitter. When mapping the social network formed by humans to an Internet-based digital representation, a peer-to-peer (P2P) architecture suggests itself by being the homomorphic choice. Recruiting the participants for data management and communication processing is not only beneficial from the privacy perspective but can even be advantageous for the scalability of a system. The absence of a single point of data aggregation removes the most powerful attacker from the list of possible adversaries. Besides privacy aspects, decentralization also addresses the issue of different objectives of the central-SNS provider (advertising) and the user (availability of data even if not attractive for advertising-mining any longer) as pointed out by Schwarzkopf et al. (2011). Another major advantage of P2P systems is the user's physical ownership of data, with benefits such as avoiding censorship, higher resilience with respect to network outages, portability, independence from provider changes to terms-of-service. In this paper, however, we focus on the privacy implications.

A. Architectures

Several architectures for social networks have been proposed, ranging from the currently predominant logically centralized, such as Facebook or Google+, to completely decentralized P2P networks (*PeerSoN*, *Safebook*, *Persona*), with hybrid architectures in between, e. g. *Diaspora* (joindiaspora.com) using a pod



(a) All communication is relayed or mediated by the central provider.

(b) Besides direct communication between two peers, several nodes can be involved.

Figure 1: Differences in architecture and communication flows of a) centralized and b) decentralized Social Network Services.

model that hosts several users or *SuperNova* (Sharma and Datta, 2011) with a distinction between super peers and normal users. A unified storage model includes options such as users hosting their own content on a local web server, home routers or individually in cloud storage or virtual machines , and replicated content on P2P clients.

To replace the functionality formerly provided by the central provider or to improve the service performance, network members are recruited for administrative services. We refer to users that take these roles as **dedicated nodes**. One example are storage nodes, that replicate the content of a user to ensure availability even when she is offline (cf. Rzadca et al., 2010, for the problem of choosing optimal replica placement in a decentralized P2P network). Furthermore users may act as communication relay nodes in order to obfuscate communication endpoints (e. g. in *Safebook*). We do not concentrate on a specific DOSN implementation, but use an abstract model as in previous work (Greschbach, et al., 2012). The DOSN is assumed to be based on a P2P overlay that connects the members of the network where content is self-hosted by the user with replicas on storage nodes (to increase availability and resilience). Figure 1 sketches the architectural differences between a centralized and decentralized system as well as abstract communication flows between nodes in these networks.

B. Encryption

In a distributed system, without a central access authorization for user content, confidentiality and integrity of user data can be ensured by cryptographic means. This way, one can prevent unauthorized access by appropriate key management and encryption of content, rather than by having to trust the SNS provider to manage access control and to not misuse data. Several encryption schemes tailored for DOSN application have been analyzed by Bodriagov and Buchegger (2011).

In the following we assume encryption to be in place for published content and traffic between members of the network. We further assume that the storage services rely on the content encryption for access control and do not perform authentication themselves so that the ciphertext objects are available to all members of the network.

C. Friends

Social relationships between users in a SNS can be of a symmetric kind (mutual friendships only, such as on Facebook) or an asymmetric kind (following someone, such as on Google+ or Twitter). We consider the social layer of a SNS as a directed graph, where vertices represent users and edges friendship relations between the users. An edge from user A to user B denotes a friendship relation, where A follows B (for example A having put B into a circle on Google+). We denote a direct edge between two users as degree-one friendship and speak of a degree-n friendship of user A to user B if there is a direct edge from user A to another user, having a degree-$(n-1)$ friendship with user B.

Information about social relations is mainly used for four reasons. First, for active access control to determine who can access a user's data (active access rights by outbound social links, i. e. following someone in Google+). Second, for passive access control to determine whose data a user can access (passive access rights by inbound social links, i. e. to be in other's circles on Google+). Third, for information preservation to organize (filter, group, prioritize) the presentation of other users' data. And fourth, for friendship announcements to show one's social relations to other users.

For the friend adversary model we require that there exists a (degree-n) friendship between the target user and the attacker. The most common case is $n = 1$ but if extended access rights also apply for example for a degree-two friendship (friend-of-a-friend), an adversary can even profit from an indirect friendship so that we explicitly include this case.

IV. Friend Adversary Model Analysis

In current SNS, the central provider is omnipotent. Here, we investigate the decentralized case. As outlined before, storage, access right management, retrieval, and other administrative tasks of the service may be delegated to the DOSN users themselves. This entails, that the members of a network are put in the position to abuse these roles. Both random sniffers and friends now have additional capabilities when compared to the centralized SNS. We focus on the friends as attackers, as they not only subsume the notion of sniffers or rogue dedicated nodes, but complement the information gained from traffic analysis with background knowledge about a specific user.

We want to point out, that we do not assume that friends by default turn into adversaries in a DOSN system. Our concern is that in the case where there is an adversary, this one becomes more powerful when being able to exploit its social ties with the target user and that this strategy can be facilitated by certain implementation issues of a DOSN system. In the following we discuss different properties of the friend adversary, namely attractive attack targets, available information sources, possible inferences from collected data, the impacts of attacks that misuse this knowledge and incentives to conduct an attack.

A. Attack Targets

Privacy preservation concerns both user data content and information about that data. We distinguish the following types of potential targets for privacy breaches: user-generated content, metadata thereof, information about social relations, and user behavior. **User content** comprises all active contributions of a user to the system, such as profile information, posts (text, picture, video, link), comments to posts, liking posts, tagging, status updates, asynchronous messaging (private messages, wall posts), synchronous messaging (chat, calls). Special functionality such as events can be seen as compositions of the mentioned primitives. **Metadata** concerns sensitive personal information that does not stem from the content of the published data but from static properties of that data (such as size or structure) or information generated while managing the data objects. This attack target is relevant even if the content is properly encrypted: The ciphertext representation of a stored object still gives away an approximate size of the content and the modification of the content will be reflected in a change of the ciphertext which can be observed by an adversary even if she is not able to decrypt the content. **Social relation information** mainly consists of the in- and outbound links of the user's node in the social graph but includes even statistical information about the interaction with certain peers that may indicate qualitative aspects of the relationships. And, finally, **behavioral data** is about usage patterns of the service that are reflected in communication flows or logs of dedicated nodes.

All four categories contain information items that are attractive targets for privacy invading attacks. Information of these types either potentially contribute to user profiling or represent critical personal knowledge that can be used against a user in various ways.

B. Information Gathering

There are several information sources that are available to an attacker without any social link to the user. Different aspects of the proposed DOSN architectures open up attack vectors to adversaries that simply sniff or crawl the network. We note that in a centralized SNS these attack vectors were only available to the central provider. We distinguish between three different strategies for information gathering. First, spreading out stored data in the network might allow an attacker to infer privacy invading information from the metadata even when the content itself is encrypted. Second, observing network communications can give away social relationship information and behavioral data about the users at the connection endpoints. Third, the encryption scheme that is used to implement the content access right management may leak further social relationship information and behavioral data. All three mentioned problems are aggravated when a friend adversary is in the position to exploit its social relationship status, so we will focus on this case. Some proposed DOSN implementations store encrypted user content preferably on friend nodes (e. g. *Safebook*). In these systems a friend adversary does have extended access to ciphertext representations of the user's data that might not be intended to her. While she cannot decrypt the content the adversary is put into the position to monitor access requests from other users as well as modifications of the object. Furthermore the ciphertext itself might give away some information. The size of the encrypted object can be an indicator for the content type (such as text-post, image, video) or – if the type is already known by other means – for statistical information such as a word-count estimate or the length of a video.

For traffic analysis attacks the friend adversary may have the advantage of being situated closer to the target user that is to be monitored. Moreover the friend adversary may operate a relay node for the user and therefore achieves an even better coverage of the user's communication. Even if the system employs encrypted connections the attacker can learn behavioral information (usage statistics, online times) and social relationship information (communication peers, contact frequency) from the intercepted traffic.

Depending on the encryption scheme used, encryption headers, access control lists or key-management operations might give away personal information about a user. The size of an encryption header that is stored together with the ciphertext object can be analyzed in order to infer for example the number of people allowed to access the content. If keys or encryption headers are reused, objects with the same content audience can be identified. Adding or removing a friend may furthermore trigger key distribution operations that can be observed by an attacker.

Besides these more passive information collection methods, a friend adversary can mount powerful active attacks to retrieve sensitive personal information. If storage nodes for content replication are chosen by indicators such as online availability or free storage space, an attacker might actively offer these resources to acquire more content objects for metadata analysis. Furthermore, a friend adversary is more likely to have suitable knowledge at it's disposal for social engineering attacks.

C. Information Mining

After collecting information about an attack target, an adversary will try to infer as much knowledge as possible from the data that might be useful for an attack against the user. The advantage of a friend adversary for this task lies in potential **background knowledge** about the user. This background knowledge might be acquired outside the social network system and therefore represents a unique advantage compared to other adversary models. These two data sources – collected information and background knowledge – can be combined to interpret existing data or generate new knowledge.

Suppose for example that the adversary has collected coarse grained location information about the user, e. g. by observing it's IP-address. If the adversary additionally holds background knowledge about likely whereabouts (workplace, home, friends, favorite cafés, etc.) the precise geographic location of the user can be inferred with high probability. In some cases it might even be sensitive information to infer that the user is not situated at a certain location at a certain time.

Another example is the observation of communication peers. In a DOSN architecture communication between two peers is more likely to be carried out directly rather than mediated by a central party. By sniffing the network an adversary might therefore be able to observe times, frequency and types of connections to certain nodes, the user is interacting with. Only a friend adversary might, however, be able to infer sensitive information from this kind of statistics by combining them with background knowledge about the user's relationships to these communication peers.

D. Attack Impact and Incentives

Another characteristic that distinguishes a friend adversary from a socially unrelated attacker is the type of impact, an attack can have on the user.

Some information might be more sensitive when disclosed to a friend than when disclosed to a stranger. Furthermore a friend attacker can also exploit extended credibility of its social status when using information against the user (e. g. claims of a stranger or a company might appear less trustworthy than accusations made by a friend).

Finally there are more user-specific incentives for a friend adversary to exploit data. In the classical setting of a centralized service provider, the main driving force for privacy invasions is an economic one. The collection of personal information and preferences of the SNS users allows for targeted advertising which is more profitable the more detailed the digital dossiers are. For a friend adversary the motivation might be of a personal, non-financial kind. Active and targeted attacks – which are costly but also more powerful than only passive crawling or sniffing attacks – are more likely in this model.

V. Discussion

In this section we discuss the novelty aspects of the friend adversary model in a DOSN system compared to a friend adversary in a centralized SNS and the central provider adversary. Furthermore we sketch possible countermeasures to mitigate these new threats to user privacy. These aspects and implications of the friend adversary model are not constraint to the field of SNS but apply to other P2P systems that have trust-based access control schemes as well.

A. New Challenges

The capabilities and available means of an adversary differ depending on the architecture and the adversary model as summarized in Table 1.

In a centralized architecture all data is collected in a single place and available to the service provider. The central provider adversary therefore does have access to all content, independently to whom it is intended, complete social relationship information and system data.

If we compare the friend adversary in a centralized SNS with a friend adversary in a decentralized system the differences are apparent. Given the security measures of a centralized system are properly implemented, a friend adversary does only have access to content that was explicitly intended to her as well as background knowledge about the user. In the decentralized setting, data is spread out over all users of the system. This allows a friend adversary to use data from the system level such as metadata about content, behavioral data and additional social relationship information as outlined in section IV.B. Compared to the central provider adversary a friend adversary in a DOSN system has the advantage of possible background knowledge that was acquired outside the system. Although not having a complete picture about the users activities in the system this might allow inferences that are not possible for the central provider. Note that there is no single point of data aggregation in the decentralized system, so the central provider adversary becomes irrelevant for that architecture.

architecture:	centralized		decentralized
adversary model:	provider	friend	friend
shared user content	✓	✓	✓
private user content	✓		
system- and metadata	✓		✓
background knowledge		✓	✓

Table 1: Adversary Information Sources

B. Protection Measures

Data that on its own does not contain privacy critical information might in combination with other data sources contribute to privacy invading inferences. Thus, it is important to limit the leakage even of seemingly insignificant information to all members of a DOSN system. Besides content confidentiality and integrity, a privacy-preserving social network has to minimize the inferences an attacker can make about a user and their social relations from information *about* the user data (e.g. by ciphertext or traffic analysis). This includes static and dynamic information about access instances to data (frequency, time, from which address, at which storage), the data itself (size, structure, updates, timestamps), access rights to the data (encryption headers, access control lists), traffic between users (frequency, size, addresses), etc.

To address inferences from the ciphertext representations of stored objects, padding the content before encryption (appending random data to obfuscate the exact size) or splitting objects in uniform block sizes (and hiding the connection between them) are possible countermeasures.

To minimize privacy leakages caused by the implementation of access control mechanisms, specialized cryptographic techniques are one possible solution. Attribute-based encryption, used for example in the *Persona* project (Baden et al., 2009), allows to define groups that can be reused by other users without them learning the explicit recipient list (and therefore enabling friend-of-a-friend access schemes). The attribute access structure stored with the object, however, might still allow inferences about the audience, e. g. by the attribute names carrying semantic meaning. Bodriagov and Buchegger (2011) propose to use broadcast encryption with hidden access structures that do not reveal anything about the audience of the content.

Considering the threat risk to be the product of occurrence probability and damage impact one can come up with two fairly opposed strategies to minimize it: either reducing the probability of a privacy breach by leveraging trust relations to friends, e. g. store content at friends' nodes, or minimizing the damage impact by *not* entrusting content to friends but personally unrelated nodes of the social network.

VI. Conclusion

Removing the threat of a centralized SNS provider also entails losing some protection offered by the walled-garden model with clear insiders, outsiders, and gatekeepers. Privacy-preserving communication applications such as but not limited to social networks, need to cope with increased capabilities of other attackers once the omnipotent adversary is removed. In the case of social networks, it is not possible to decrease the level of knowledge a friend on the online social network has of a user based on out-of-system observations and gossip. It thus becomes even more important to provide the technological mechanisms for privacy protection within the system itself to avoid a combination of inferences from metadata and friend knowledge.

After thoroughly investigating the threat and adversary model, the challenge lies in designing appropriate countermeasures to limit possible inferences from system information. Just one example is to deploy encryption mechanisms that do not reveal who else has access to a given piece of data. A systematic and careful approach is needed to preserve user privacy and refrain from adding new privacy threats when introducing solutions for previous privacy concerns.

References:

AIELLO, L. M. AND G. RUFFO. 2010. Secure and flexible framework for decentralized social network services. *In: 2010 8th IEEE International Conference on Pervasive Computing and Communications (PERCOM) Workshops*. pp.594–599.

BADEN, R. AND A. BENDER AND N. SPRING AND B. BHATTACHARJEE AND D. STARIN. 2009. Persona: An Online Social Network with User-Defined Privacy. *In: ACM SIGCOMM Computer Communication Review*. 39(4), pp.135.

BODRIAGOV, O. AND S. BUCHEGGER. 2011. Encryption for Peer-to-Peer Social Networks *In: Workshop on Security and Privacy of Social Networks at the IEEE International Conference on Social Computing*. Boston: IEEE, pp.1302–1309.

BUCHEGGER, S. AND D. SCHIÖBERG AND L.-H. VU AND A. DATTA. 2009. PeerSoN: P2P social networking: early experiences and insights. *In: Proceedings of the Second ACM EuroSys Workshop on Social Network Systems 2009*. New York: ACM Press, pp.46–52.

CUTILLO, L. A. AND R. MOLVA AND T. STRUFE. 2009. Safebook: A privacy-preserving online social network leveraging on real-life trust. *In: IEEE Communications Magazine*. 47(12), pp.94–101.

GRESCHBACH, B. AND G. KREITZ AND S. BUCHEGGER. 2012. The Devil is in the Metadata – New Privacy Challenges in Distributed Online Social Networks. *In: Fourth International Workshop on Security and Social Networking (SeSoc12)*. pp.339–345.

GROSS, R. AND A. ACQUISTI 2005. Information revelation and privacy in online social networks. *In: Proceedings of the 2005 ACM workshop on Privacy in the electronic society (WPES '05)*. pp.71–80.

KRISHNAMURTHY, B. AND C. E. WILLS. 2010. On the leakage of personally identifiable information via online social networks. *In: ACM SIGCOMM Computer Communication Review*. 40(1), pp.112–117.

PAUL, T. AND S. BUCHEGGER AND T. STRUFE. 2011. Decentralized Social Networking Services. *In:* L. Salgarelli, G. Bianchi and N. Blefari-Melazzi, eds. *Trustworthy Internet*. Milan: Springer Milan, pp.187–199.

RZADCA, K. AND A. DATTA AND S. BUCHEGGER. 2010. Replica Placement in P2P Storage: Complexity and Game Theoretic Analyses. *In: 2010 IEEE 30th International Conference on Distributed Computing Systems*. pp.509–609.

SCHWARZKOPF, M. AND A. MADHAVAPEDDY AND T. HONG AND R. MORTIER. 2011. Personal Containers: Yurts for Digital Nomads. *Available from: perscon.net*.

SHARMA, R. AND A. DATTA. 2011. SuperNova: Super-peers Based Architecture for Decentralized Online Social Networks. *In: Fourth International Conference on Communication Systems and Networks*.