Safebook: a Privacy Preserving Online Social Network Leveraging on Real-Life Trust

Leucio Antonio Cutillo* Refik Molva* Thorsten Strufe † *Institut Eurécom
Sophia Antipolis
France
{cutillo, molva}@eurecom.fr
†TU Darmstadt
Darmstadt
Germany
strufe@cs.tu-darmstadt.de

Abstract—**On-line social network applications severely suffer from various security and privacy exposures. This paper suggests a new approach to tackle these security and privacy problems with a special emphasis on the privacy of users with respect to the application provider in addition to the defense against intruders or malicious users. In order to assure users' privacy in the face of potential privacy violations by the provider, the suggested approach adopts a decentralized architecture relying on the cooperation among a number of independent parties that are also the users of the on-line social network application. The second strong point of the suggested approach is to capitalize on the trust relationships that are part of social networks in real life in order to cope with the problem of building trusted and privacy-preserving mechanisms as part of the on-line application. The combination of these design principles is Safebook, a decentralized and privacy preserving on-line social network application. Based on the two design principles, namely, decentralization and exploiting real-life trust, various mechanisms for privacy and security are integrated into Safebook in order to provide data storage and data management functions that preserve users' privacy, data integrity and availability. Preliminary evaluation of Safebook shows that a realistic compromise between privacy and performance is feasible.**

I. Introduction

Social Networking Services (SNS), like *facebook*, *LinkedIn*, or *orkut*, are a predominant service on the web, today. Catering for a broad range of users of all ages, and a vast difference in social, educational, and national background, they allow even users with limited technical skills to publish personal information and to communicate with ease. In general, the Online Social Networks (OSN) that are stored for this purpose are digital representations of a subset of the relations that their participants, the registered persons or institutions, entertain in the physical world. Spanning all participating parties through their relationships, they model the social network as a graph. However, the popularity and broad acceptance of social networking services as platforms for messaging and socialising attracts not only faithful users, who are trying to add value to the community, but parties with rather adverse interests, be they commercial or plain malicious, as well.

The main motivation for members to join an OSN, to create a profile, and to use the different applications offered by the service, is the possibility to easily share information with selected contacts or the public, for either *professional*, or *personal* purposes. In the first case, the OSN is used as a facility geared towards career management or business goals, hence SNS with a more serious image, like XING or LinkedIn, are chosen. As members in this case are aware of the professional impact of the OSN, they usually pay attention to the content of the data they publish about themselves and others. In the case of a more private use, they share more personal information like contact data, personal pictures, or videos. Other members in the shared pictures can be marked ("*tagged*"), and links to their respective profiles are created automatically.

The core application used by the members of the SNS is the creation and maintenance of their contact lists, which describe the members' milieux and maps them into the digital OSN graph.

Through informing members automatically on profile changes of their contacts, the SNS thus helps users to stay up to date with news of their contacts and very often the popularity of users is measured in the number of contacts their profile links to.

These properties of the services have led to the definition of boyd and Ellison [1], according to which *Social Network Sites* or *Online Social Network Services* are:

> *web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.*

This definition, however, leaves aside an additional service that becomes apparent when observing the use of SNS: the communication of members through direct, sometimes instant message exchange, or annotation of profiles (*giving comments* or *recommendations*). Additionally, SNS typically enable a wealth of third-party applications featuring advanced interactions between members ranging from simple *"poking"* of another member or the support for *interest groups* for a

This work in parts has been supported by the SOCIALNETS project, grant no 217141, funded by the EC FP7-ICT-2007-8.2 for Pervasive Adaptation.

common topic to *"likeness"* testing with other members and the exchange of virtual *gifts*.

Storage, maintenance and access to the OSN and their services are offered by commercial providers, like Facebook Inc.[1], LinkedIn Corp.[2], Google Inc.[3], XING AG[4], and the likes.

Analyzing the OSN with respect to their security properties and the privacy of their users, some obvious threats become apparent. Generally, a wealth of personal data on the participants is stored at the providers, especially in the case of OSN targeting non-professional purposes.

This data is either visible to the public, or, if the user is aware of privacy issues and able to use the settings of the respective SNS, to a somewhat selected group of other members. As profiles are attributed to presumably known persons from the real world, they are implicitly valued with the same trust as the assumed owner of the profile. Furthermore, any actions and interactions coupled to a profile are again attributed to the assumed owner of this profile, as well. Different studies have shown that the participants clearly represent the weak link for security in OSN and that they are vulnerable to several types of social engineering attacks[5,6,] [2], [3]. This partially is caused by a lack of awareness to the consequences of simple and presumably private actions, like accepting contact requests, tagging pictures, or acts of communication like commenting on profiles or leaving wall posts. However, the usability of privacy controls offered by the SNS, and finally and most importantly inherent assumptions about other participants and trust in other profiles, which are actually a desired characteristic, certainly add to the problem.

However, analyzing the privacy problems in current OSN it becomes apparent, that even if all participants were aware and competent in the use of SNS, and even if a concise set of privacy measures were deployed, the OSN would still be exposed to potential privacy violations by the omniscient service provider: the complete data, directly or indirectly supplied by all participants, is collected and stored permanently at the databases of the providing company, which potentially becomes a big brother capable of exploiting this data in many ways that can violate the privacy of individual users or user groups. The importance of this privacy exposure is underlined by the market capitalization of these providers, which ranges from 580 million US$ (acquisition of myspace through the news corp. in 2005) to 15 billion US$ (Facebook Inc, according to the investment of Microsoft in 2007)[4].

In consequence, we consider the protection of private data in OSN a pressing topic, which current providers are not likely to address. In this paper we suggest a SNS called Safebook[7] that is specifically designed to prevent privacy violations by intruders, malicious users or OSN providers alike. Safebook is mainly characterized by a decentralized architecture relying on the cooperation among the peers, in order to prevent potential privacy violations due to centralized control. In addition to the description of Safebook, this paper presents

- a multi-layered model of social networking services, and
- a security analysis of threats and attacks in online social networking.

Section II states the security objectives for online social networks. Section III analyzes the security requirements of current social networking services and section IV presents Safebook, our new approach to a privacy preserving SNS. Finally, we conclude with a summary and an outlook in section VI.

II. Security Objectives in OSN

In the context of online social networks, we generally identify three main security objectives, *privacy*, *integrity*, and *availability*, which come in slightly different flavors than in traditional systems.

A. Privacy

In accordance to previous studies [5], [6], we assume the protection of the user's privacy to be the main objective for SNS. Privacy does not only encompass the protection of personal information, which users publish at their profiles, presumably accessible by their contacts only. But additionally, communication privacy has to be met. Hence, no other than directly addressed or explicitly trusted parties may have the possibility to trace, which parties are communicating. Furthermore, details of messages have to be hidden, so only the requesting and responding parties should know one another's identity and the content of the request. Finally, disclosure of information about a third party to some member that is not explicitly trusted by the third party, without the consent of the latter, has to be prevented. In summary, privacy calls for the possibility to hide any information about any user, even to the extent of hiding their participation in the OSN in the first place. Moreover privacy has to be met by default, i.e., all information on all users and their actions has to be hidden from any other party internal or external to the system, unless explicitly disclosed by the users themselves.

Requiring explicit disclosure directly leads to the need for *access control*. Access to information on a user may only be granted by the user directly, and the access control has to be as fine grained as the profile and each attribute has to be separately manageable.

B. Integrity

As part of integrity, the user's identity and their data must be protected against unauthorized modification and tampering. In addition to conventional modification detection and message authentication, integrity in the context of OSN has to be extended: parties in an OSN are not arbitrary devices, but real, unambiguously identifiable persons. The creation of personae – bogus accounts, cloned accounts, or other types of impersonation – in traditional SNS is easy to achieve. However, users have a strong inherent trust in OSN, and it

[1] www.facebook.com
[2] www.linkedin.com
[3] www.orkut.com
[4] www.xing.com
[5] http://www.sophos.com/pressoffice/news/articles/2007/08/facebook.html
[6] http://www.blackhat.com/html/bh-usa-08/bh-usa-08-archive.html
[7] www.safebook.us

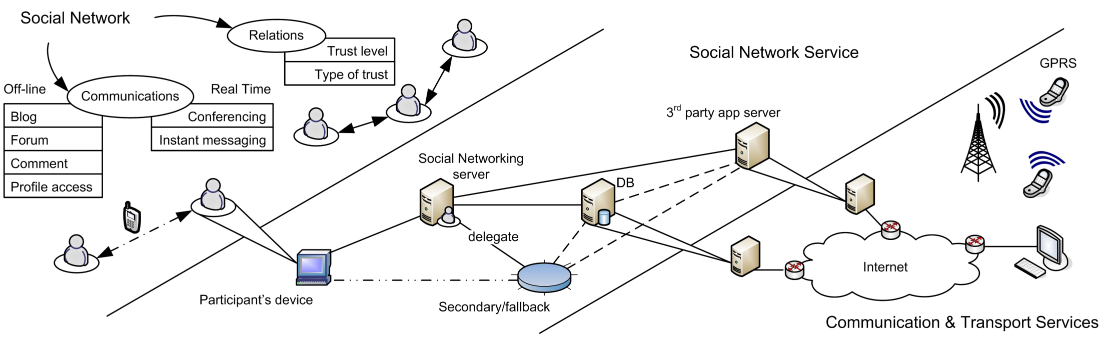

Fig. 1. OSN levels: three architectual layers of social networking services.

has been shown that this combination may lead to a new kind of vulnerabilities[5,6,][3]. In consequence, the authentication has to assure the existence of real persons behind registered OSN members. Identity checks do not necessarily have to be performed by a centralized service, however, all identification services have to be trusted by all participants.

C. Availability

Since some SNS are used as professional tools to aid their members' business or careers, data published by the users has to be continuously available. Availability of user profiles in consequence is required as a basic feature, even though considering recreational use, the availability of some content may not seem a stringent requirement. In online social networks, this availability specifically has to include a robustness against censorship, and the seizure or hijacking of names and other key words. Apart from availability of data access, availability has to be assured along with the message exchange among members.

III. Security Analysis of OSN

First of all, we shall sketch a model for SNS, to get an overview on the aim and possible implementation schemes of SNS.

Social Networking Services can be divided into three different levels (cmp. Fig. 1):

- a ***Social Network*** (SN) level: the digital representation of members and their relationships;
- a ***Social Networking Service*** (SNS) level: the application infrastructure, managed by the SNS provider;
- a ***Communication and Transport*** (CT) level: communication and transport services as provided by the network.

The SN level provides each member with a set of functions corresponding to social interactions in the real life, like finding friends, accessing profiles, commenting, and the like.

To implement these functions, the SN level relies on the SNS level. This second level includes the infrastructure managed by the SNS provider, together with basic services to create the SN service, such as web-access, storage, and communication. Common strategies to enhance availability for these are redundancy and delegation: both for organizational reasons or if a server faces failures or other inabilities to provide a service, it may delegate requests to secondary servers. Data storage and retrieval, indexing of the content, management of access permissions to data, and node join or leave, are implemented in a centralized or decentralized, distributed fashion on the SNS level.

The SNS level on the other hand relies on the transport and internetworking protocols and infrastructures, implemented by the CT level.

Based on this architecture of OSN, we define an attacker as one of:

- a malicious member on the SN level;
- a malicious service provider on the SNS level;
- a party that has and misuses access to the infrastructure at the CT level (an eavesdropper with a local, or a malicious ISP with possibly even a global view).

Other than these inside attackers, that primarily seem to be legitimate participants in the system but act in a malicious way in some cases, there may be external attackers, or *intruders*. An intruder can perpetrate attacks at one or more of the SNS levels.

After defining the different levels of SNS, we shall characterize major attacks on SNS.

Privacy The protection of a member's identity is one of the key aspects that still have to be addressed in current OSN. In ***Identity Theft***, e.g., a malicious member or service provider acquires the credentials of authorized users and acts on their behalf with full access to this profile, relations and communication traces. Due to the inherent trust in other profiles, plain ***Impersonation*** by creation of a ***Clone*** of the targeted profile[8] may suffice to be able to establish trust relationsships with parties on a victim's contact list by simply sending new friendship requests. ***Profile Porting*** attacks, in which the attacker creates a profile under the victim's identity in an OSN where the victim is not present, are more difficult to detect. However, with most existing accounts being unprotected, profile porting poses a valid threat. The collection of existing data is the basis of ***Profiling*** attacks, data aggregation that leads an attacker to the possibility to guess the value of a potentially huge set of, usually disclosed, properties, such as the victim's social security number, income bracket, potential interest in some product, etc. They additionally supply potential attackers with the knowledge needed for ***Secondary Data Collection***, as from the data published

[8] http://www.nature.com/news/2009/090423/full/news.2009.398.html

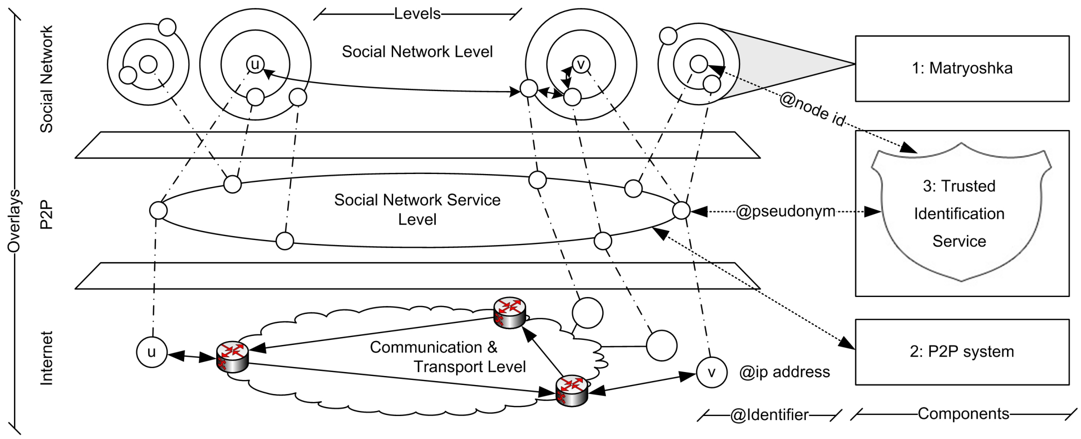

Fig. 2. Safebook overlays (left) and main components (right)

on OSN they may easily be able to guess the social security number (or, e.g., the "Foedselsnummer" in Norway), which often acts as a key to access personal information from a wide range of different sources. Matching the profiles of a person in both a professional and a more informal OSN for analysis and comparison of the content published in both is another obvious and frequent type of secondary data collection.

Moving from data to communication privacy, a series of other threats arises. A malicious SNS provider or, to some extent, a malicious member with the appropriate set of privileges, can be able to perform ***Communication Tracking*** and reveal who is talking to whom. The problem becomes relevant, and much more difficult to solve, at the CT level with an omniscient ISP.

Another series of attacks on privacy is ***Profile Harvesting***, in which an attacker, a malicious participant or SNS provider, gathers data on the participants on a large scale for purposes that the victims have not considered, intended, or foreseen. More sophisticated harvesting comes in the form of ***Image Retrieval***, possibly even in association with automated ***Face Recognition*** algorithms for further profiling.

Integrity The abovementioned impersonation threats are due to a basic shortcoming: none of the current major OSN is able (nor interested, in many cases) to ensure that a profile is associated to a single, real person. ***Faked profiles*** are a common phenomenon resulting from this shortcoming, as well as clones, or ported profiles. Such impersonation paves the way for ***Sybil attacks***, which aim at creating fake identities, as well as ***Defamation*** and ***Ballot Stuffing*** attacks that aim at forging the reputation for a person using the system, or at disrupting digital reputation systems.

Availability Several types of ***Denial of Service***, e.g., to cover a victims profile or selected data, or to disrupt the possibility to communicate with a victim, are possible in SNS: A centralized OSN obviously is vulnerable to ***Censorship*** through the SNS provider. However, distributed SNS, which are implemented as decentralized, possibly P2P, systems, or which follow other types of service delegation, may be vulnerable to a series of attacks that known are from these domains. ***Black Holes***, ***Selective Forwarding***, and ***Misrouting*** are serious threats in this case.

Table I gives an overview of the relationship between the stated attacks and the involved security objectives. Some attacks breach several objectives, but still primarily focus, or mainly exploit a vulnerability to one of these objectives, in which case they are attributed to only this objective. Attacks that are not mainly related to a single security objective, like e.g. collusions, have to be encountered by a number of measures regarding different objectives and they are hence attributed to more than one objective.

In conclusion it becomes apparent that current SNS still are vulnerable to different attacks on all the three different levels by either insiders (legitimate parties) or outsiders (intruders). In the following section we will describe Safebook, a new approach to decentralize SNS, to convey this approach as an alternative solution to open vulnerabilities, that are unlikely to be fixed by current SNS providers.

IV. Decentralized Online Social Networking

Some of the security and privacy exposures analysed in III could be addressed through the enhancement of existing OSN applications, by integrating various security and privacy mechanisms. However, the privacy of the users' data is at risk due to the central storage and management and hence threatened by potentially malicious service providers or unintended access following short-sighted publication[9], security breaches, or plain misconfiguration of the OSN. It inherently cannot be assured with centralized, server-based architectures, on which all existing OSN rely. Peer-to-peer architectures seem to offer a suitable alternative to the centralized approach as the basis for a decentralized OSN avoiding the all-knowing service provider. As a major drawback, P2P systems suffer from a lack of a-priori trust, thus creating the need for cooperation incentives. We thus suggest a decentralized OSN based on a peer-to-peer architecture whereby basic security and privacy problems as well as the lack of a priori trust and incentives are addressed by leveraging on real life trust between the users,

[9] http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006

Security Objectives **Attacks**	Privacy	Integrity	Availability
Id Theft	x	x	x
Profile Cloning	x	x	
Profile Porting	x	x	
Secondary Data Collection	x		
Profiling	x		
Communication Tracking	x		
Face Recognition	x		
Image Retrieval	x		
Harvesting	x		
Fake profiles		x	
Sybil	x	x	x
Ballot Stuffing		x	
Defamation		x	
Censorship			x
Collusion	x	x	x

TABLE I
ATTACKS VS SECURITY OBJECTIVES IN ONLINE SOCIAL NETWORKS, PRIMARILY AFFECTED OBJECTIVES ARE HIGHLIGHTED

such that services like data storage or profile data routing are performed by peers that trust one another in the social network.

A. *Safebook: Security based on Real-Life Trust*

Safebook consists of a three-tier architecture with a direct mapping of layers to the OSN levels depicted in Fig. 2 as follows:

- the user-centered Social Network layer implementing the SN level of the OSN;
- the Peer-to-Peer substrate implementing the SNS services;
- the Internet, representing the CT level.

Each party in Safebook is thus represented by a node that is viewed as a host node in the Internet, a peer node in the P2P overlay, and a member in the SN layer.

The nodes in Safebook form two types of overlays:

- a set of ***Matryoshkas***, concentric structures in the SN layer providing data storage and communication privacy created around each node;
- a P2P substrate, providing lookup services.

In addition to these nodes, Safebook also features a ***Trusted Identification Service*** (TIS), providing each node unambiguous identifiers: the ***Node Identifier*** for the SN level and a ***Pseudonym***.

Each Safebook component plays an essential role since it implements a particular set of countermeasures against the threats presented in section III.

Matryoshka Matryoshkas are concentric rings of nodes built around each member's node in order to provide trusted data storage, profile data retrieval and communication obfuscation through indirection. Each matryoshka thus protects the node in its center, the ***core***, which on the SN layer is addressed by its Node Identifier. The nodes in the matryoshka are connected through radial paths on which messages recursively can be relayed from the outermost shell to the core and vice versa. All paths are based on trust relationships akin to the social network, thus each hop connects a pair of nodes belonging to users linked by a trust relationship in real life. The innermost and the outermost shell of a matryoshka have a specific role: the innermost shell is composed of direct contacts of the core, and each of them stores the core's data in an encrypted form. Hence they are called the ***mirrors***. Every node in the outermost shell acts as a gateway for all the data requests addressed to the core, and is thus called ***entrypoint*** (cmp. Fig. 3). All requests to a core are addressed using its node identifier. Real time communication is responded to by the core itself, any kind of off-line communication can be served by one of its mirrors as well. While the number of mirrors and entrypoints in each path is fixed, the number of nodes between them is variable, thus leading to paths with variable length on the same Matryoshka.

Peer-to-peer system In order to provide a location service to find entrypoints for a user's Matryoshka, the nodes create a P2P substrate. Currently, this substrate resembles a KAD[10] and the pseudonyms are used as identifiers for the DHT. The searchable and registered keys are the hashed properties of the participating members and their node identifiers. Unlike the path across a Matryoshka, the communication through the P2P layer does not rely on trusted links. However, using pseudonyms still protects the members from privacy violations based on node identification and tracing through the untrusted P2P links.

Trusted Identification Service (TIS) The TIS assures that each Safebook user gets at most one unique identifier in each category of identifiers. Based on an out of band identification

[10] xlattice.sourceforge.net/components/protocol/kademlia/specs.html

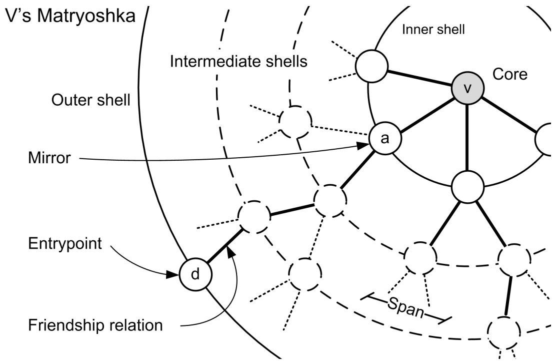

Fig. 3. Matryoshka structure

procedure, the TIS grants each user a unique pair of a node identifier and a pseudonym, computed as the result of a keyed hash function on the set of properties that uniquely identify a party in real life, such as full name, birth date, birth place and so on. Even if at a first glance a centralized trusted third party service such as the TIS seems to contrast the purpose of decentralization as pursued by Safebook, the TIS, even though being a centralized service provider, does not pose as a privacy threat, as it cannot trace users, their messages, nor can it peek into their private data. Moreover, the TIS can be implemented in a distributed and off-line fashion.

B. Operations

Safebook implements different OSN operations:

- account creation;
- data publication;
- data retrieval;
- contact request and acceptance;
- message management.

1) Account creation: In order to join Safebook, a new member $\mathcal{V}$ has to be invited by one of its real life friends $\mathcal{A}$ that must be already a registered user. $\mathcal{V}$'s account is then created in the two steps of the identity- and the matryoshka creation.

Identity creation After $\mathcal{A}$'s invitation, $\mathcal{V}$ provides the TIS with its identity property set $name_v$, together with a proof of owning it. This credential request contains also the public keys $\mathcal{P}_v^+$ and $\mathcal{I}_v^+$ belonging to two keypairs $\mathcal{P}$ and $\mathcal{I}$, which are generated by $\mathcal{V}$ itself. The TIS then computes the node identifier of $\mathcal{V}$ and its pseudonym by applying two different keyed hash functions with two different unknown master keys to $name_v$. At this point, the TIS sends $\mathcal{V}$ back its pseudonym P_v and node id I_v together with the certificates $Cert(P_v, \mathcal{P}_v^+)$ and $Cert(I_v, \mathcal{I}_v^+)$ associating the peer and member identifiers of $\mathcal{V}$ to its public keys $\mathcal{P}_v^+$ and $\mathcal{I}_v^+$ respectively. The pseudonym keypair $\mathcal{P}$ is used to guarantee integrity and confidentiality to all the messages exchanged in Safebook, as in each hop every message is signed using the sender's pseudonym private key and encrypted using the receiver's pseudonym public key, while the node id keypair $\mathcal{I}$ is used to guarantee the same properties to end-to-end communication between members

It becomes evident that, even if a valid member $\mathcal{V}$ repeats the account creation operation multiple times, it will always receive the same pseudonym and node identifier, since they are a function of $\mathcal{V}$'s identity itself. Moreover, $\mathcal{V}$ cannot claim the ownership of an identity that is not its own, since it wouldn't be able to prove this fact. The identity proof is an out of band process that relies on real life mechanisms to ascertain the identity of a potential member, such as face to face meeting between a user and the representation of the TIS, or relying on existing tamper-proof schemes such as passport, id card, etc. According to this fact, sybil and impersonation attacks are not possible, as $\mathcal{V}$ cannot manipulate its node id nor its pseudonym.

Once $\mathcal{V}$ gets its identifiers, it can join the P2P system by using $\mathcal{A}$ as a bootstrapping node and start the matryoshka creation process.

Matryoshka creation $\mathcal{V}$ has only $\mathcal{A}$ as a contact to start with, so it sends $\mathcal{A}$ a request for path creation containing the DHT lookup keys it wants to register, a ttl, and the number of members $\mathcal{A}$ should forwards the request to, hereafter called the ***span*** factor. $\mathcal{A}$ then selects between its friends a number $span$ of next hops and forwards them this registration message. This process is recursively done until the ttl expires: the receiving node $\mathcal{D}$ registers the lookup key in the P2P system together with its reference $@d$ and starts acting as an entrypoint for $\mathcal{V}$.

Matryoshkas provide for privacy based on hop-by-hop trust, as all nodes in each Matryoshka are only aware of their direct neighbors.

As soon as $\mathcal{V}$ has created its matryoshka, it can publish its profile (cmp. Fig. 4).

2) Data Publication: The data managed in SNS can be generalized to:

- profile information;
- contact relations;
- messages.

The profile information is the part of the data each user intends to publish. To guarantee a fine grained access control, it is organized in atomic attributes for which particular access policies can be set. Contact relations represent member's real life relations and can be seen as the friend list of the user. As the strength of a relation is not the same for all links [7], in Safebook each user associates a particular trust level to each of its contacts. This level is used to select closely related contacts that primarily will store the published data. Finally, personal messages or comments on profiles can be exchanged between members. In case of comments, the receiver has the right to publish or discard them.

To guarantee privacy, data in Safebook can be private, protected or public: in the first case the data is not published, in the second case it is published and encrypted, in the third case it's published without encryption. All the published data of a member $\mathcal{V}$ is replicated to its mirrors, the nodes in the innermost shell of $\mathcal{V}$'s matryoshka.

3) Data retrieval: The lookup of $\mathcal{V}$'s data through the member $\mathcal{U}$, starts with a recursive query in the P2P system: according to the DHT structure, the node responsible for the lookup key responds with the entrypoint list building $\mathcal{V}$'s outer shell. Consequently, $\mathcal{U}$ can request that one of $\mathcal{V}$'s entrypoints forwards the request through $\mathcal{V}$'s matryoshka, until a mirror is

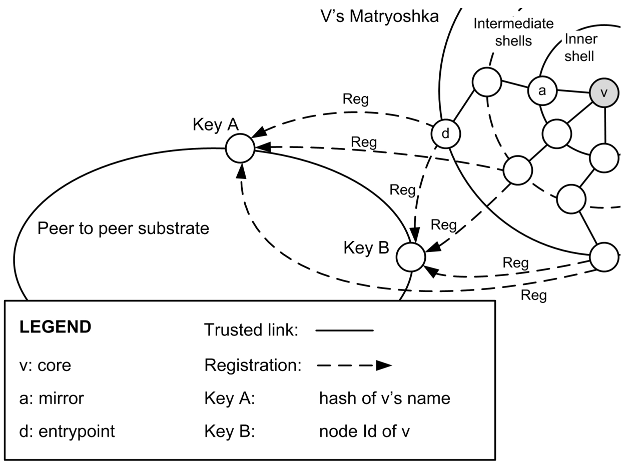

Fig. 4. Entrypoint registration in the P2P substrate

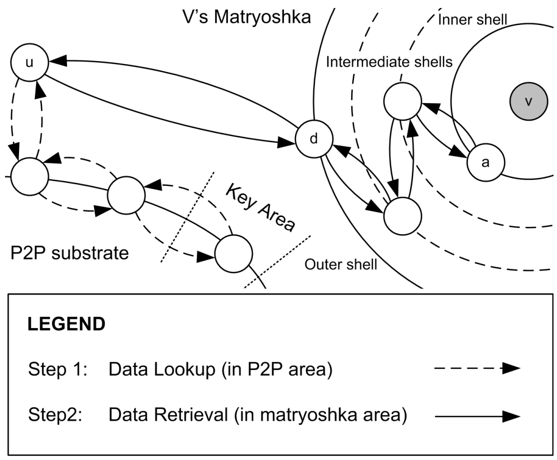

Fig. 5. Data lookup and retrieval.

reached. $\mathcal{V}$'s encrypted data then reaches $\mathcal{U}$ through the inverse path (cmp. Fig. 5).

The protocol of Safebook uses recursion to hide the source of requests. Additionally, the addressing and routing, both for P2P lookup and for data retrieval using the Matryoshkas, are based on the pseudonyms of nodes. Attackers in consequence have no means to identify a source of a request for some content, as there is no way to distinguish between generated and forwarded requests. Since the mapping between the pseudonym of a node and its identifier is only known to the TIS and direct connections ("friends") in the Matryoshka, which are trusted by the node, no private information can be derived from it either. Finally, communication tracking is not possible, as a malicious node would always have to be the first hop for all requests to the Matryoshka of a certain node in order to be able to link the pseudonym of the sender to its real identity.

A preliminary feasibility study conducted with a previous and less performant approach [8] showed that data retrieval performs well, even though the messages are forwarded along multiple hops in the overlay.

4) Contact request and acceptance: A member $\mathcal{U}$ that wants to add another member $\mathcal{V}$ to its contact list sends a contact request message following the same steps as in the data request case. Assuming $\mathcal{V}$ accepts $\mathcal{U}$ as a new contact, $\mathcal{V}$ associates to $\mathcal{U}$ a certain trust level (known by $\mathcal{V}$ and nobody else) and sends it back an opportune key that will enable $\mathcal{U}$ to decrypt the selected parts of $\mathcal{V}$'s published encrypted data.

5) Message management: Off-line messaging, such as wall posts, recommendations, and other annotations to a profile, is implemented using the steps of retrieving some members data, decrypting the shared parts, annotating some content, and sending this data back, signed with the key bound to the annotator's node identifier and encrypted with the public key bound to the receivers node identifier. On reception of this updated message the receiving mirror advertises it to the other mirrors and to the adressed node that finally can choose to sign and republish, or to discard it.

Real-time messages, like chats, are forwarded to and handled by the core solely and responded to with an error message if the core is off-line.

Hence, in Safebook, only members with appropriate privileges can access and update the profiles of other members. Entity and data authentication are provided through common signatures and encryption schemes.

V. Related work

While a series of studies [6],[2], [3] has investigated privacy and security exposures of current OSNs, several other articles suggest solutions to these exposures in various directions combining cryptography with advanced distributed computing techniques.

The approach in NOYB [9] mitigates the existing problems by cryptographic means thanks to the application of substitutions according to secret dictionaries. Public profiles, which still may be stored in centralized OSN, are thus made useless to anybody lacking access to these dictionaries. Whereas some of the contents of the profiles are protected, this is not the case for the relations between users, as expressed by contact lists or message exchange.

Yeung et al [10] propose using the existing World Wide Web Friend-Of-A-Friend representation of people and their relations as an OSN. Conventional content and friendship relations are stored in the user's personal space hosted by a server, the choice of which is left at the discretion of the users. While access control for user data can be efficiently assured based on articulated policies, the system does not protect the identity of the users.

Persona [11] offers flexible and fine-grained access control for user data by combining attribute-based encryption with traditional public-key cryptography. Users are identified by public keys they exchange out of band while creating OSN links, while data confidentiality and privacy is assured through encryption. Users have to trust a firefox extension to interact with Persona and can also create multiple identities.

The related work closest to Safebook is probably PeerSon [12]. PeerSon achieves decentralization thanks to an external P2P system, OpenDHT, and assures access control through encryption. Whereas it represents a fully distributed OSN, PeerSon leverages on an untrusted P2P system and thus offers a weaker privacy protection than Safebook.

Although not designed originally for the purpose of social networking, darknets and related P2P systems [13], [14], [15] aim at anonymous communication through hop-by-hop encryption among trusted users, as in Safebook. Unfortunately, such systems suffer both from delays that could be prohibitive for OSN.

VI. Conclusion and Future Work

This paper outlined a new approach for the design of online social networks that addresses privacy problems akin to existing social network applications. Potential access to the private data of users, such as profiles and contact lists, and possible misuse of such information by the providers of social networking services is viewed as the highest privacy exposure. In order to assure users' privacy in the face of such potential exposure, the suggested approach adopts a decentralized architecture relying on the cooperation among a number of independent parties that are also the users of the online social network. The second strong point of the suggested approach is to capitalize on the trust relationships that are part of social networks in real life in order to cope with the problem of building trusted and privacy-preserving mechanisms as part of the SNS. The result of these design principles is Safebook, a decentralized and privacy preserving SNS. Various mechanisms for privacy and security are integrated into Safebook in order to provide data storage and data management functions that preserve privacy, data integrity and availability. The current design and prototyping of Safebook raise an interesting trade-off between privacy and performance. While increasing the number of hops through trusted links increases privacy, it severely affects lookup and communication delays. A preliminary evaluation of Safebook shows that a realistic compromise between privacy and performance is feasible. Fine tuning of the performance models and simulation results also help determining critical design parameters such as obfuscation layers and data replication factors. Furthermore, the underpinnings of Safebook can serve as a model to tackle various problems that were left unsolved in the area of secure communications. Thus, a decentralized approach relying on social links can shed new light on hard problems of the past such as anonymous communications, secure routing or cooperation enforcement in self-organizing systems.

L. Antonio Cutillo is a Ph.D. student at EURECOM in Sophia Antipolis, France. He received his M.S. in Computer Engineering from the Polytechnic of Turin in 2008, his Diplôme d'Ingénieur en Systèmes de Communication from EURECOM and his Master Research in Image and Geometry for Multimedia and Life Modelization from TELECOM ParisTech in 2007. He actually works at the Networking and Security Department dealing with security and privacy concerns in distributed systems under the supervision of Prof. Refik Molva.

Refik Molva is a professor at EURECOM. His research interests are the design and evaluation of protocols for security and privacy in self-organizing systems. He was program chair or general chair for security conferences such as ESORICS, RAID, SecureComm, IEEE ICC and security workshops. He is an area editor for the Computer Networks Journal, Computer Communications Magazine and the Pervasive and Mobile Computing Journal. He worked in the Zurich Research Laboratory of IBM as one of the key designers of the KryptoKnight security system.

Thorsten strufe is professor for Peer-to-Peer Networks at Technische Universität Darmstadt. His research interests lie in the areas of decentralized distributed systems and security, with an emphasis on network analysis and the construction of resilient systems. Recently, he has focused on studying privacy and security in online social networks. and possibilities to provide social networking services through P2P technologies. Previously, he took a post as senior researcher at EURECOM, and at TU Ilmenau, working on resilient networking technologies.

References

[1] d. m. boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship." *Journal of Computer-Mediated Communication*, vol. 13(1), 2008.

[2] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Communications of the ACM*, vol. 50, no. 10, pp. 94–100, 2007.

[3] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks," in *18th Intl. World Wide Web Conference (WWW'09)*, 2009.

[4] "Modelling The Real Market Value Of Social Networks," http://www.techcrunch.com/2008/06/23/modeling-the-real-market-value-of-social-networks/, 2008.

[5] R. Gross and A. Acquisti, "Information revelation and privacy in online social networks," in *ACM Workshop on Privacy in the Electronic Society*, 2005, pp. 71 – 80.

[6] danah m. boyd , "Facebook's privacy trainwreck," *Convergence: The International Journal of Research into New Media Technologies*, vol. 14(1), pp. 13 – 20, 2008.

[7] W. X. Zhou, D. Sornette, R. A. Hill, and R. I. M. Dunbar, "Discrete hierarchical organization of social group sizes," *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1561, pp. 439–444, 2005.

[8] L.-A. Cutillo, R. Molva, and T. Strufe, "Safebook: Feasibility of Transitive Cooperation for Privacy on a Decentralized Social Network," in *World of Wireless, Mobile and Multimedia Networks*, 2009.

[9] S. Guha, K. Tang, and P. Francis, "NOYB: Privacy in Online Social Networks," in *Online Social Networks*, 2008, pp. 49–54.

[10] C. M. A. Yeung, I. Liccardi, K. Lu, O. Seneviratne, and T. Berners-Lee, "Decentralization: The Future of Online Social Networking," in *Future of Social Networking*, 2009.

[11] R. Baden, A. Bender, D. Starin, N. Spring, and B. Bhattacharjee, "Persona: An online social network with user-defined privacy," in *ACM SIGCOMM*, Barcelona, Spain, August 2009.

[12] S. Buchegger, D. Schiöberg, L. H. Vu, and A. Datta, "PeerSoN: P2P Social Networking," in *Social Network Systems*, 2009.

[13] M. Rogers and S. Bhatti, "How to Disappear Completely: A Survey of Private Peer-to-Peer Networks," 2007.

[14] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, "Freenet: A Distributed Anonymous Information Storage and Retrieval System," in *Design Issues in Anonymity and Unobservability*, 2000, pp. 46 – 66.

[15] K. Bennett and C. Grotho, "GAP - Practical Anonymous Networking," in *Privacy Enhancing Technologies*, 2003, pp. 141–160.