Modeling Role-Based Privacy in Social Networking Services

Gábor György Gulyás, Róbert Schulcz, Sándor Imre
Mobile Innovation Centre
Budapest University of Technology and Economics
Budapest, Hungary
{gulyasg, schulcz, imre}@hit.bme.hu

Abstract—**As social networking services are getting more and more common, the need for privacy enhancing options, sophisticated identity management and anonymity emerges. In this paper the authors propose using Role-Based Privacy as a response for these needs and introduce a novel model called Nexus-Identity Network that is capable of describing services extended with such functionality. The concerned principles of Role-Based Privacy are conferred in the paper and criteria are presented for anonymity. Conforming to the criteria the authors suggest storing the profiles of different identities in a tree hierarchy in a user-friendly manner. The analysis of anonymity shows that the network has a structure that can be easily interpreted similarly to graphs representing connections in regular social networks. The ease of profile management and network visualization are advantages of the Nexus-Identity Model which can make a social networking service privacy- and user-friendly as well.**

Anonymity; privacy; identity management; social networks

I. Introduction

On the World Wide Web today, there are numerous community based services and yet new services appear every day – in most cases labeled as Web 2.0 services. Similarly to these, one can find other services on the Internet using social networking mechanisms basing their values on social connections, user participation and contribution.

Access and information management issues arise as user interactions are frequent in these services. Users need to control who may access their profiles or part of their profiles (prevention of profiling and protecting sensitive information), prohibit unwanted messaging activities (spamming) or hide parts of their social web. In other cases one might want to manage multiple but unlinkable personae, for instance manage separately a political identity from the professional one.

Since there are several user groups and individuals in social networking services in one's neighborhood that should be handled separately, the authors propose to adopt the principle of data minimization by using the technique of Role-Based Privacy (RBP). This Identity Management (IDM) technique is already used in business and federated applications, like in the PRIME Project [1]. However, in these types of services the users' goals are to strengthen their privacy against business services providers, instead of other users: direct user-user interactions are not common, if present at all. There is some other work related to the latter topic; authors in [2] propose using RBP in these scenarios.

In this paper the analysis of a novel model is presented that is based on Social Networks (SN) extended with RBP functionality. As SNs are usually represented with undirected or unidirectional graphs, the model needs to represent multiple user identities, anonymous connections and communities.

The next section summarizes the requirements for social networks using RBP and for the model as well, followed by Section III, in which the criteria for anonymity is conferred in details. In Section IV the novel model called Nexus-Identity Networks (NIN) is introduced and questions regarding anonymity and network structure are discussed. Finally, the work presented in this paper is concluded in Section V.

II. Motivation and Requirements

There are known privacy-related issues in social networks [3], although there is a need for enhanced user-centric privacy control. Some problems are mentioned in the introductory section, but there are others for example enhanced profile management or anonymity under specific circumstances. These issues can be solved by binding profile management to user-defined roles and letting the user manage her neighbors' access accordingly. These roles may also be shared in communities, rooms or to other user groups within the social networking service.

The following requirements are derived by analyzing the most important concerns of RBP in social networks. All principles listed here should be noted, but since the model introduced in Section IV describes the static state of the social networking services (similarly to social network graphs), also requires these principles.

Granularity and perspicuity of control. Managing the roles set for user groups and individuals should be easy and unambiguous. For example some roles may be related to a specific business relationship and used for a lot of transactions and some others may be used just for a single person or transaction. Accordingly, this means that the model should describe how users can control access to their profiles, visibility of related events, possible interactions with them. Profiles may include presence information as well.

Community types. Usually several community types exist within a given service as communities may have several important properties regarding user privacy. Types may vary accordingly to these properties. For instance a community can be local (no global listing) and invitation based or anonymity may be supported.

Control issues. A set of operations should be defined in every community type for role management. For instance an operation may be used to ignore messages from someone, and another to introduce an identity change in a community, supporting anonymity by granting unlinkability of the old and the new identities.

Triggers. Triggers help users manage their privacy. Dependent upon the service, triggers on geo-locations, system events or time tables may be used to manage privacy settings and roles.

Anonymity. Anonymous presence and the chance to act anonymously should be made available at least in some of the communities. Two levels of anonymity should be considered: unlinkability of identities and total anonymity (no identifiers). The principles of anonymity criteria are conferred in the following section.

III. Anonymity Criteria

Anonymity needs to be guaranteed on the network and application level separately as different attack types should be prevented and different data should be protected against leakage, forgery or modification. The authors call the principle of this separation the *inner-outer world paradigm*, since participants inside the service and outsiders should be separated: latter should not be able to interfere with application level protocols.

This paradigm is in accordance with previous works in the field of privacy, for instance with the abstraction of the Nymity Slider [4]. The slider tells about the amount of personal information revealed in a given transaction scaling from unlinkable anonymity (nothing is shared) to verinymity (exact identification) and has an interesting property: the less information is shared the easier it is to increase the level of possible identification. Analogously, the less information is shared on the network level the more options a user may have on the application level.

The network architecture should include an anonymizer protocol to protect the network level of the communication and applications protocols – meaning identity management protocols in this case – should be designed to allow anonymity.

A. Network Architecture

As numerous social networking services are centralized the proposal presented here is also for centralized services, consequently the network architecture should consist of two main parts: application servers and an anonymizing service, a special network running an anonymizer protocol. Anonymizing services hide network level information such as IP addresses, port numbers and also protect against traffic analysis and the observation of messages sent through the network.

The concept of anonymizing networks was first defined by David Chaum when MIX networks were introduced [5]. A MIX network is a set of MIX nodes, allowing its users to send messages anonymously. Besides the protocol allows recipients to send reply messages to the anonymous sender. A MIX node achieves anonymity by removing the correlation between its input and output messages. Using a simple MIX node had a serious drawback: a single node may not be trusted. Hence, MIX cascades are used instead a single node.

Today, there are several anonymizing networks providing sender anonymity [6]; and many general-purpose networks are acceptable for centralized social networking services like TOR or I2P.

B. Anonymity Criteria

The following properties are required for anonymity in an Identity Management System (IMS) by applying the technique of RBP. The criterion of *unobservability*, *unlinkability* and *pseudonymity* are the main concerns for *anonymity*. The criteria are similar for business and federated RBP [7].

Unobservability. An attacker, as an outsider, may not gain information by observing the communication. This means unobservability of message content or the impossibility of timing analysis of events (such as timings of messages).

Unlinkability. Usually this property means unlinkability of two or more entities, such as arbitrary many messages, some messages and a user or a sender and a receiver. However, in an IMS the unlinkability of identities should also be concerned: an attacker should not be able to link seemingly independent identities to a certain user or create identity sets belonging to an unknown user (grouping identities).

Pseudonymity. A pseudonymous user is recognized by an identifier number or string, which cannot be related to any personal (or any globally unique) information.

Based on the previous criteria, three levels of anonymity are distinguished: *pseudonymous identification*, *unlinkable pseudonymity*, *total anonymity*, which are defined as follows.

Pseudonymous identification: a user is identified by a unique global pseudonym, but the identifier is not linkable to the user as a person. This is certainly a level of anonymity, but in IDM this should not be interpreted as real anonymity. In this case the user may have arbitrary many identities; however, these identities are linkable by the global pseudonymous identifier.

Unlinkable pseudonymity: a user may have arbitrary many pseudonymous identities, but these are unlinkable due to the lack of the attached global identifier (identities might not be linked to a certain user). An event is only linked to the originating identity. The user should be reckoned as anonymous in this case since there is no unequivocal relation between identities and users; an identity change executed properly guarantees the unlinkability of the prior and the subsequent one.

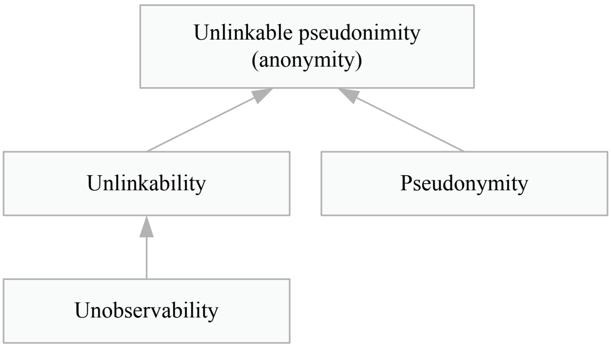

Figure 1. Relationship of the anonymity criteria properties.

Total anonymity: no pseudonymous identifiers should be linked to the user and all actions committed by the user should also be anonymous. Total anonymity should be at least an option.

The relationship of the properties is illustrated on Fig. 1. An IMS that allows anonymity and uses the technique of RBP should provide unlinkable pseudonymity at least or total anonymity optionally.

In some cases redesigning the protocols is not possible and filter mechanisms should be applied (usually transparently by proxies). For instance in the case of web privacy it is not possible to redesign application protocols, which indicates the need for applying privacy enhancing extensions [8].

IV. Nexus-Identity Networks

Nexus-Identity Networks is a graph based model, to describe the *static state of connections in social networks* with RBP functionality; however, the principles behind the model also propose a how-to for introducing RBP in social network based services.

A graph of a regular social network is a set of users as nodes and a set of edges describing connections between identities. For instance such an undirected graph may represent map of social connections between individuals (who knows who) on a social networking site or a unidirectional connection map may represent connections between mobile telephony users accordingly to the phonebook entries.

In NINs another node type is introduced called *nexus* as a generalization of communities. In their offline lives people often associate roles with groups, which also motivated the introduction of communities. Also, users usually contact these groups accordingly to the role they play. Therefore, in our model users interact in nexuses which includes profile management issues (sharing and receiving profile information), sending and receiving messages, events, and nexuses are even used to handle access control issues. NIN networks omit identity-identity connections and they are understood as graphs describing connections between identities and nexuses only.

The introduction of nexuses has advantages regarding the modeling concerns of privacy: a nexus allows modeling any of the three levels of anonymity previously conferred. Unfortunately, due to the lack of communities total anonymity can not be interpreted in social networks.

Terms *identity* and *profile* are discerned in our work. Let us define profiles as data sets describing an identity, an identity as the virtual person herself – however, a real person might have numerous identities. Profiles may include the presence information of the identity as well. A user may have several identities. In case these identities are unlinkable then the user is able to achieve anonymity as unlinkable pseudonymity.

Issues related to anonymity are conferred in the following section with some examples for Nexus-Identity Networks in Section IV-C.

A. Model Properties

NINs are defined formally with a graph model similarly to social networks. The latter may be described as an undirected graph such as $G_{SN} = (V, E)$, where V is the set of users and E is the set of connections. Let a NIN be a directed graph such as $G_{NIN} = (V_i \bigcup V_N, E_{i \to N} \bigcup E_{N \to i})$, where V_i are the set of identities, V_N are set of nexuses and $E_{i \to N} \bigcup E_{N \to i}$ are connections between. The basic structure of the networks is illustrated on Fig. 2.

The outgoing and incoming edges of an identity model several functions. The outgoing edge models profile information shared with other identities through the nexus, but also models the possibility of sending messages and events (or even controlling the nexus). Accordingly the incoming edge represents subscription to the nexus, which means that the identity receives messages (or events) and shared profiles through the nexus.

These edges may have certain states as being hidden or anonymous – a pair of these edges may be handled independently. Therefore there are three possible types of connections: only presence information is distributed, only listener status or both. In some systems restrictions may apply in specific nexus types. The authors suggest that if a user is only present as a listener then anonymous presence should be noted.

For instance, in a service independent scenario a user

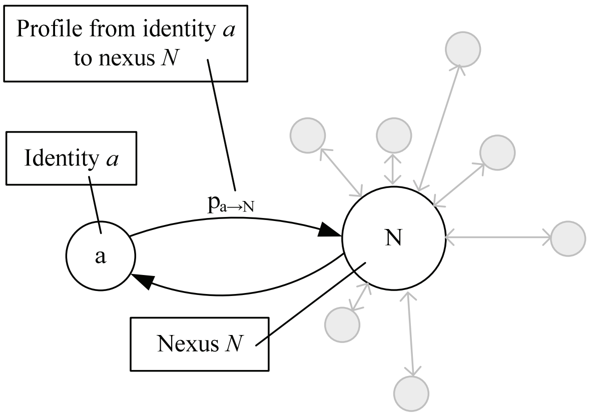

Figure 2. The basic structure of Nexus-Identity Networks.

may open a channel to share her profile and presence information with others. This can be modeled as a nexus to which the user shares a profile and others may subscribe, but share no profiles. The subscription process may be controlled or limited by the user. Some more examples are discussed in Section IV-C.

The achievable level of anonymity is determined by the identifier the nexus requires. Some may require a globally unique identifier (e.g., registration ID), but others may only require a locally unique identifier. The latter should be unique within a set of nexuses which set is defined as a domain for unique identifiers. In a NIN there is a global domain and there may be several local domains. Some nexuses can be excluded from all domains.

B. Profile Management

Realizing profile management in the proper way is a key concern for usability. The authors suggest ordering profiles in a tree hierarchy as shown on Fig. 3. There is inheritance in the structure as profiles on lower levels inherit their values (and settings) from higher levels unless the user sets it otherwise. This profile hierarchy can also be applied for any service types described by a NIN model. This model is familiar with the spanning tree of roles mentioned in [7].

The achieved level of privacy and anonymity depends on the profile settings at the current profile. As profiles may contain locally or globally known identifiers, all levels of anonymity are available. Total anonymity is achieved if no identifier is set in the profile; otherwise if the nexus is in a local domain then unlinkable pseudonymity is achieved.

C. Examples for Modeling Services

Many services can be modeled with NINs. The modeling process involves three levels of types: the parts of the original NIN model are meta-types from which the types of the modeled system are derived (second level) and finally from these types the instances for the current static state of the service are constructed. This section gives a few examples on deriving service types in different services.

Social Networks. Usually there are two separate profiles in social networks: public and private, both should be interpreted as different nexuses for which the owner shares a profile and others may subscribe to. A subscription is interpreted as the subscriber knows the owner – in some services mutual subscriptions are required. Fan groups, communities, forums are also modeled as nexuses including system specific restrictions as subscriptions are handled together with profile sharing, meaning that a connection is interpreted as being a member of the community.

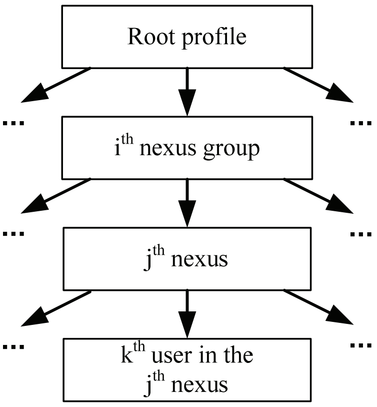

Figure 3. Profile hierarchy.

Instant Messaging (IM), Chat services. In these services the network is allocated from contact lists (buddy lists) or name lists in rooms, conferences. In some services all features may be present and can be used jointly to create the connection map. Contact lists consist of user groups that are represented as nexuses for sharing profile information and others' profile information is gathered by similar means.

Conferences and rooms are similar; however, there is a relevant difference: a global identifier is required and displayed in conferences. Therefore only rooms are capable of offering pseudonymous unlinkability. Both communities are modeled as nexuses (presence distribution and subscription are handled together), but in some cases anonymous presence and contribution is available in rooms meaning that the choice for total anonymity can be offered.

For instance rooms can be divided into several local domains. In this case the profile hierarchy would contain a path such as the following: root profile, 1st local domain for rooms, ith room, jth user in the ith room.

Mobile telephony may also be modeled as NINs. A user may have several identities by using several Subscriber Identity Modules (SIM) at the same time and social connections are represented by contact entries in the SIM card or the mobile phones.

Implementing RBP in these services requires an anonymous sanction system as anonymity prohibits prevention against malicious activities in some cases. Private credentials are also important in business and federated RBP services [7].

Private credentials may be used in social networking services as a private passport describing entries of reputation which values are managed by the service when malicious acts (e.g., spamming) are committed. This passport should be unavailable for reading for any users; however, it should be possible to formulate constraints regarding certain actions (e.g., joining a room may require a clean passport with no entries of spamming).

D. Analysis of Anonymity

Total anonymity and unlinkable pseudonymity are the two desirable levels of anonymity to reach in social networking services with RBP – in Nexus-Identity Networks it is possible to achieve both levels. The main concern is unlinkability of an identity with others (grouping identities) or a unique global user identifier.

In case of *total anonymity* the user connected to the nexus can not be recognized by legal means and only the correlation of behavior (including writing style, message content, topic, etc.) or action timings can be used to compromise the user's anonymity. Anonymous observers can not be compromised without committing any actions. Distributing the number of anonymous subscribers in the nexus may also help to observe anonymous presence and activity.

The case of *unlinkable pseudonymity* is more complex and there are several sub cases. The level of anonymity can degrade for an identity *a* due to linkability issues.

Sub case 1. There is a globally unique identifier related to identity *a*. Obviously, instead of unlinkable pseudonymity this is pseudonymous identification.

Sub case 2. There are some other identities with the same pseudonym, but all identities are within the same local domain. Since both identities *a* and *a'* have the same locally unique identifiers they must be the same.

Sub case 3. Otherwise, the linkability of identity *a* is uncertain with a user with or other identities, additionally if these identities are in different domains, it is not even necessary for their identifiers to differ. The linkability of identities needs to be proven by correlating profiles, behavior, timing of actions or based on other information.

For instance there may be some other clues suggesting linkability: two identities with differing identifiers can not join the same nexus are likely to belong to the same user (reason: a user with several identities in the same nexus would be a disturbing phenomenon, hence should be disallowed).

E. Network Structure

The linkability of user profiles correlates with the structure of the network which differs greatly from the structure of social networks. Since two linkable profiles are represented as a single node, the more linkable profiles can be found the fewer distinct parts there are in the structure of the network. Therefore the visible structure (for users) of the network is always a disconnected graph, in which each subgraph is generated by the different domains. Additionally there may be nexuses outside all domains – there are no name uniqueness restrictions within these nexuses.

A possible layout is shown on Fig. 4. Nexuses are noted as empty circles, identities as dots. Connections are undirected as they may represent any types of connections (anonymous connections are not shown as they do not affect network structure).

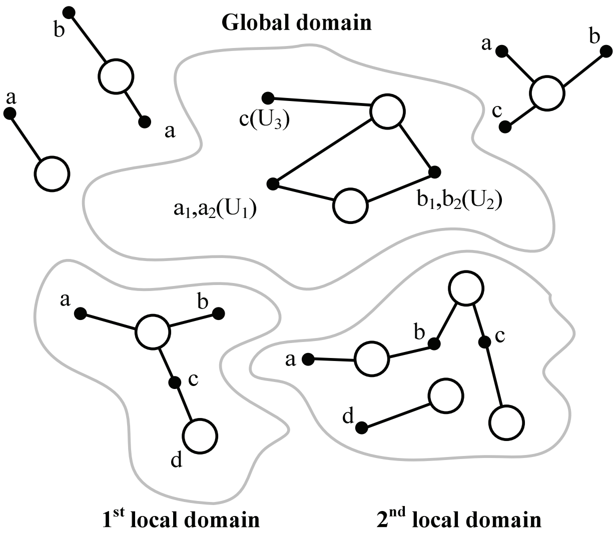

Figure 4. Network structure example.

In this case the NIN consists of a global and two local domains. In the global domain of nexuses of groups a globally unique identifier is required; the user identifier is denoted next to the list of used identifiers. Local domains are on the bottom of the figure; as the second local domain shows subgraphs induced by domains are not always connected.

The structure of the network suggests that it is not possible to detect network wide paths between individuals based on visible information. This is desirable to detect whether identities are cooperating to correlate linkability of identities based on visible information (profile, events, etc.) received through nexuses.

There are some other properties that can affect the structure. In some services users may reveal their global identifier in local domains creating bridges between subgraphs. Requiring that identifiers in the global domain need to be unique throughout the whole network also creates bridges between the subgraphs of the network.

V. Conclusion and Future Work

The principles behind Nexus-Identity Networks support greatly introducing Role-Based Privacy into social networking services; furthermore the model has a similar representation to social networks, although the network structure may be unconnected due to supporting anonymity.

In the authors' opinion this way of visualization can help to understand the current situation from the point of view of anonymity and the concept of the profile hierarchy can be considered helpful also as it makes identity management issues easier to see through. Applying both features can make a social networking service more user-friendly and also enhance privacy protection, but making interviews and surveys regarding this matter is considered as future research assignment.

However, as the model is new these are some other areas that need further research, such as focusing future work on the analytic study of anonymity. Additionally, Nexus-Identity Networks are incapable of describing possible operations for managing access control and profile issues, temporal changes within the structure which research tasks should be also assigned as future work.

Acknowledgment

This paper was made in the frame of Mobile Innovation Centre's integrated project Nr. 3.1 supported by the National Office for Research and Technology (Mobile 01/2004 contract).

The authors would also like to thank for the support of the HSN Lab (High Speed Networks Laboratory, website can be found at http://hsnlab.tmit.bme.hu).

REFERENCES

[1] R. Leenes, J. Schallaböck, and M. Hansen, “PRIME white paper (V3)”, 15/05/2008.

[2] J. Hakkila, and I. Kansala, “Role based privacy applied to context-aware mobile applications”, 2004 IEEE International Conference on Systems, Man and Cybernetics, Volume: 6, 10-13/10/2004, pp. 5467- 5472.

[3] M. Chew, D. Balfanz, and B. Laurie, “(Under)mining Privacy in Social Networks”, W2SP 2008, Oakland, California, USA, 22/05/2008.

[4] I. Goldberg, “A Pseudonymous Communications Infrastructure for the Internet”, PhD Thesis. 2000.

[5] D. Chaum, “Untraceable electronic mail, return addresses, and digital pseudonyms”, The Communications of the ACM 24, February, 1981, pp. 84-88.

[6] A. Jones, “Anonymous communication on the internet”, WWW@10 Conference, Terre Haute, Indiana, USA, 01/10/2004.

[7] M. Hansen, A. Schwartz, and A. Cooper, "Privacy and Identity Management", IEEE Security and Privacy, vol. 6, no. 2, Mar/Apr, 2008, pp. 38-45.

[8] G. Gulyás, R. Schulcz, and S. Imre, “Comprehensive Analysis of Web Privacy and Anonymous Web Browsers: Are Next Generation Services Based on Collaborative Filtering?”, Joint SPACE and TIME International Workshops 2008, Trondheim, Norway, 17/06/2008.