On the Security of Social Networks

Literature review on ReBAC policies and Sybil defenses

Alceste Scalas <alceste.scalas@unica.it>

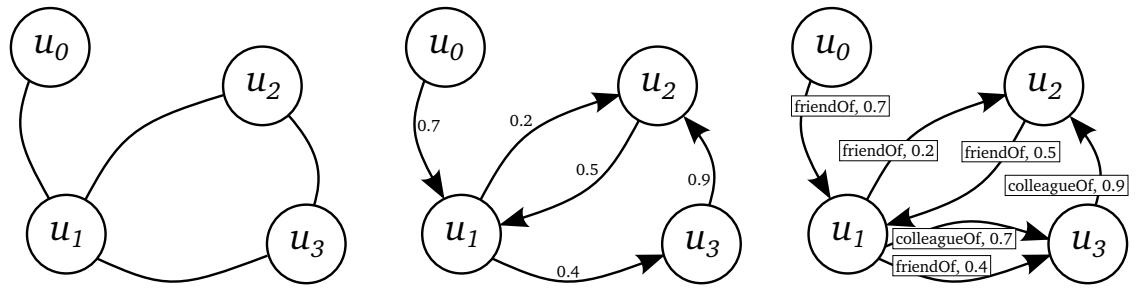
University of Cagliari — Department of Mathematics and Informatics

November 2012

Social networking is part of the lives of millions of people. Strong privacy concerns make it urgent to regulate how personal resources (e.g. messages, photos...) are made available to other people. One of the answers to this requirement are *Relationship-Based Access Control (ReBAC)* policies, which establish whether a requester is allowed to access the resources owned by another user, depending on the closeness and trust between them. This novel security model departs from the classical *Role-Based Access Control (RBAC)*, and its implications are still being investigated: ReBAC harnesses the shape of the *social graph*, which may be modeled and used in various ways; furthermore, such shape is defined by the social network users themselves, who can modify their relationships and/or create new pseudonymous profiles. Such modifications may just be aimed at deceiving ReBAC, thus increasing one's access privileges on other people's resources: this is the so-called *Sybil attack*. Is it possible to identify these malicious activities, or ensure that access policies are immune to such threats? This paper reviews the main approaches to ReBAC, with a particular focus on Sybil defenses.

1 The social graph

The literature reports several methods for modeling a social graph, depending on the required level of detail and/or conformance to existing implementations (see fig. 1.1). We adopt a notation similar to Fong 2011b: we consider *Social Network Systems (SNSs)* built upon a set of *social network profiles* $S = \{u_1, u_2, \ldots\} \subseteq$ *Sub (Social User Base)*. The set of all *social graphs* is $\mathcal{G}(S) = \langle S, E \rangle$, where E (the set of relationships/edges) may be either a subset of $S \times S$ (for modeling directed relationships, as in Fig. 1.1b and 1.1c) or a subset of $[S]^2$ (the set of all size-2 subsets of S, for modeling symmetric relationships only, as in Fig. 1.1a).



(a) *Simple graph with symmetric relationships (Facebook-style).*

(b) *Directed graph with trust levels (Orkut-style).*

(c) *Directed graph with relationship types and trust levels.*

Figure 1.1: *Different types of social graphs used in literature.*

2 Defining and implementing ReBAC

The ReBAC concept is first introduced in Gates 2007, where it is stated that *"a new paradigm of access control needs to be developed that is based on interpersonal relationships"* — thus capturing the essence of the research trends in social network security of the previous years. Gates' paper officially opens the problem of deciding how ReBAC could be defined and implemented: some of the most influential proposals, based on different types of social graphs, are summarized in the next two sections.

2.1 Authorization via trust metrics

Kruk et al. 2006 annotate each social graph edge with a *Friendship Level Metric* $\mathrm{FLM}_{\mathrm{context}}(u_1, u_2) \in [0, 1]$, representing the trust between profiles u_1 and u_2 within a given context. The resulting social graph is reported in Fig 1.1b. Using this information, *Social Networked Access Control Lists* can be defined taking into account both the distance (i.e., the length of the shortest path) and the "trust level" between the resource requester and its owner. Such trust level is the cumulative product of the FLMs on each path connecting two profiles: thus, in the example of Fig. 1.1b, u_0 may access u_2's resources either via u_1 (with trust $0.7 \times 0.2 = 0.14$) or via u_1, u_3 (with trust $0.7 \times 0.4 \times 0.9 = 0.252$).

A similar approach is adopted in Carminati, Ferrari, and Perego 2006, with the addition of *types* (labels) to social graph edges — thus allowing to express different relationships among profiles (e.g. distinguishing friends from colleagues, as shown in fig. 1.1c). This work is further refined in Carminati, Ferrari, and Perego 2007 and Carminati and Ferrari 2009, where the relationship themselves (which may be sensitive) are subject to access control, in a way that avoids their delegation to a centralized Social Network Management System. Decentralization (i.e. *Collaborative Access Control*) is obtained using cryptography, for certifying relationships and ensuring secrecy of communication between social network nodes.

2.2 Authorization via graph-theoretic policies

A more general approach is adopted in Fong, Anwar, and Zhao 2009 and Fong 2011b: their analysis is focused on *Facebook-style Social Network Systems (FSNS)*, with undirected edges representing symmetric friendship relationships (as shown in fig. 1.1a). Trust metrics are replaced by generic *policy predicates* $Sub \times Sub \times \mathcal{G}(Sub) \to \{\top, \bot\}$ establishing whether the resources of a profile (first argument) are accessible by a requester (second argument), in a given social graph (third argument). Some examples:

- $\texttt{dist}_k(u_1, u_2, G)$ holds if, in the social graph G, u_1 and u_2 are separated by at most k edges;
- $\texttt{clique}_k(u_1, u_2, G)$ holds in G if $u_1 = u_2$, or they both belong to a clique with k nodes;
- $\texttt{popularFof}_k(u_1, u_2, G)$ holds if $\texttt{dist}_2(u_1, u_2, G)$ and u_2 is "popular" (i.e., has at least k friends).

A FSNS N is formalized as a triple $\langle Sub, \mathcal{PV}, Pol \rangle$, where $\mathcal{PV}$ is the *policy vocabulary* (i.e., the list of policies that users are allowed to choose for regulating how other users can access, search, and traverse their profile) and *Pol* is a function associating each user to the policies he/she has selected. For instance, an user u_1 in a graph G may decide that his profile is accessible by all users u such that $\texttt{dist}_2(u_1, u, G)$ holds (i.e., u is a friend or a friend-of-a-friend of u_1). Several *policy combinators* are also introduced ($\wedge$, $\vee$, $\circ$, $(\!|\cdot|\!)$), allowing the conjunction, disjunction, composition and conditional application of policies.

This formal approach poses the problem of finding a suitable language for expressing the ReBAC policies themselves: while Fong 2011b uses a graph-theoretic tool (*bi-rooted graphs*), Fong 2011a and Fong and Siahaan 2011 use modal logic to express and compose binary relations which are sensitive to their context (i.e., the social graph). Such approach is later revised in Bruns et al. 2012: there, hybrid logic is proposed for expressing complex topological predicates — with the addition of syntactic constraints and a type system, aimed at restricting the policies to a *local* fragment of the logic itself (i.e., with predicates based on the relationship between the owner and requester profiles).

3 The Sybil attack

The Sybil attack (Douceur 2002) consists in privilege escalation through malicious alteration of the social graph (via creation of pseudonymous profiles or collusion with other users). This attack has been tackled by different authors, both in its original domain (peer-to-peer networks) and in other contexts, such as social networks "proper"*. The next sections review the main approaches. This paper does not survey the vast amount of counter-measures aimed at defusing Sybil attacks by increasing the cost of pseudonymous profile creation (e.g. by using CAPTCHAs).

*Interestingly, the "link spamming" attack to PageRank is a Sybil attack for web graphs (Cheng and Friedman 2005).

3.1 Detecting Sybil attacks from social network structure

Several papers try to detect Sybil attacks by analyzing their impact on the social network structure, and introducing criterions for establishing whether a node is "rogue" (thus reducing its privileges, even when ReBAC policies would otherwise have granted access). Yu et al. 2006 consider the networks arising in peer-to-peer systems and adopt the simple social graph model in fig. 1.1a, observing that *"malicious users can create many identities but few trust relationships"* — wile on the contrary, "real" social networks show rich interconnections (i.e. they are *fast-mixing*). The authors assume that pseudonymous identities cannot be avoided — but also notice that Sybil attacks create peculiar graph structures, shown in Fig. 3.1: they have a small *quotient cut*, i.e. the removal of a small number of edges (the *attack edges*) disconnects a large number of nodes (Sybil identities). Finding those attack edges would allow to isolate the Sybil sub-graphs — however, since the *Minimum Quotient Cut* problem is NP-hard, the authors propose a verifiable random walk of length w to partially explore the social graph starting from two nodes (*verifier* and *suspect*): if both nodes stay within a "honest" region, then (with high probability) their walks will also stay in the same region, and intersect. The authors find an upper bound to the probability of a walk to traverse an attack edge (depending on the length of the path, the number of attack edges and the size of the network), and provide experimental results of Sybil node detection on several synthetic social networks.

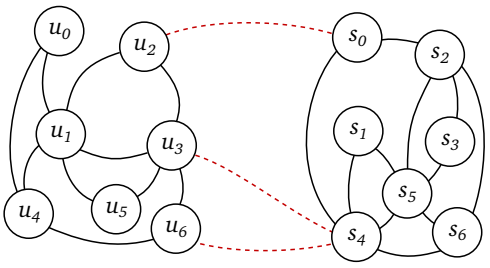


Figure 3.1: *Sybil attack: the subgraphs of legitimate nodes (left) and Sybil nodes (right) are connected by a small number of* attack edges *(dashed lines).*

These results are further improved in Yu et al. 2008, and then in Danezis and Mittal 2009 with the introduction of Bayesian inference for the detection of Sybil sub-graphs. A detailed comparison of these methods is available in Viswanath et al. 2010. However, due to their statistical nature, all these approaches are subject to false negatives and positives (thus missing real sybils, or flagging regular profiles as malicious).

3.2 Sybilproof ReBAC policies

Another approach against Sybil attacks forfeits any statistical or topological assumption on Sybil networks, and focuses in ensuring that ReBAC policies are demonstrably immune to such attack.

Cheng and Friedman 2005 consider directed social graphs with a *reputation function* $f\colon Sub \to \mathbb{R}$ which associates each profile to its trust level: they adopt a trust metrics scenario as in fig. 1.1b, and f determines how incoming edges contribute to the trust level of a node. A *Sybil attack* is successful if an user u creates an arbitrary amount of sybils $s_1, \ldots, s_n$, and among them there exists s' such that $f(s') > f(u)$. The paper aims at establishing the conditions for *sybilproof* reputation functions, and the authors find out that *symmetric* ones (i.e. invariant under graph isomorphism, thus unchanged by node renaming and based on "global" trust values) cannot be sybilproof. This leads to the necessity of *asymmetric* functions, i.e. computed with respect to some fixed node u in the graph, allowing "*reputation to propagate along paths outward*" from u (Cheng and Friedman 2005). These asymmetric reputation functions can be used *"when each user computes separately the reputations of other users with respect to themselves"* (Cheng and Friedman 2005), thus effectively representing a form of *ante litteram* ReBAC policy (similar to Kruk et al. 2006, seen in section 2.1). They are demonstrated to be sybilproof under certain monotonicity conditions.

Fong 2011b (already introduced in section 2.2) stresses that it is necessary to clearly identify the security properties that a social network (and its policies) must guarantee, since it is the only way to *"gauge the effectiveness of an access control scheme"*. For this reason, he reprises Cheng and Friedman 2005's approach, and focuses on the Sybil-proofness of ReBAC in FSNSs. The policies in his model are "asymmetric", but it may not be enough: for example, in Fig. 1.1a, u_1 may choose an access policy such as $\texttt{popularFof}_k$; in this case, u_3 or u_4 may create k pseudonymous profiles, befriend them, and thus gain the privilege for accessing u_1's resources. Hence, $\texttt{popularFof}_k$ cannot belong to the policy vocabulary of a syibilproof FSNS — but what is the criterion for establishing whether a ReBAC policy is "safe"?

Fong introduces several innovations in ReBAC research. He chooses the *Principle of Privilege Attenua-*

tion (POPA) (Denning 1976) (i.e. users can only confer privileges they already have) as the design principle for "safe" social networks. This principle, however, needs to be interpreted (and formalized) in the social network context. For this reason, Fong proposes a transition system in which the states of a FSNS N are represented by social graphs, and transition relations $G \xrightarrow{\cdot}_N G'$ correspond to befriending/defriending actions (i.e. edge addition/removal). Intuitively, a FSNS N is POPA-compliant if, whenever two profiles u_1, u_2 befriend (causing the transition $G \xrightarrow{e}_N G'$), then, for any policy p adopted by any other user $u \in N$, $p(u, u_1, G')$ must not hold unless $p(u, u_1, G)$ or $p(u, u_2, G)$ (i.e. u_1 or u_2 already had access to u's resources *before* the transition). In Fong's paper, this kind of POPA-compliance is extended to a more complete social network model, adapted to sequences (*traces*) of transitions, and it is demonstrated to be sufficient and (more surprisingly) also *necessary* for defeating Sybil attacks. Furthermore, Fong defines a static analysis to determine whether a FSNS complies to POPA, and demonstrates it to be sound and complete, i.e. *precise* (all POPA-compliant FSNS are recognized as such). Returning to the previous example, such static analysis fails if a FSNS includes $\texttt{popularFof}_k$ in its policy vocabulary. Another interesting result is that adding restrictions on social network profile search (just like Facebook does) may make the system intuitively more "secure", but also makes static POPA compliance checking fail (thus reducing the formal security guarantees): a nice confirmation of author's call against adding *"guards and mediations with no clear security goal"*.

References

Bruns, Glenn et al. (2012). "Relationship-based access control: its expression and enforcement through hybrid logic". In: CODASPY '12. San Antonio, Texas, USA: ACM. DOI: 10.1145/2133601.2133616.

Carminati, B. and E. Ferrari (2009). "Enforcing relationships privacy through collaborative access control in web-based Social Networks". In: CollaborateCom. DOI: 10.4108/ICST.COLLABORATECOM2009.8339.

Carminati, B., E. Ferrari, and A. Perego (2007). "Private Relationships in Social Networks". In: *IEEE 23rd International Conference on Data Engineering*. DOI: 10.1109/ICDEW.2007.4400987.

Carminati, Barbara, Elena Ferrari, and Andrea Perego (2006). "Rule-Based Access Control for Social Networks". In: OTM 2006. DOI: 10.1007/11915072_80.

Cheng, Alice and Eric Friedman (2005). "Sybilproof reputation mechanisms". In: P2PECON '05. ACM. DOI: 10.1145/1080192.1080202.

Danezis, George and Prateek Mittal (2009). "SybilInfer: Detecting Sybil Nodes using Social Networks". In: Network and Distributed System Security Symposium 2009. San Diego, California, USA.

Denning, Peter J. (1976). "Fault Tolerant Operating Systems". In: *ACM Comput. Surv.* 4. DOI: 10.1145/356678.356680.

Douceur, John (2002). "The Sybil Attack". In: *Peer-to-Peer Systems*. Lecture Notes in Computer Science. DOI: 10.1007/3-540-45748-8_24.

Fong, Philip, Mohd Anwar, and Zhen Zhao (2009). "A Privacy Preservation Model for Facebook-Style Social Network Systems". In: ESORICS 2009. DOI: 10.1007/978-3-642-04444-1_19.

Fong, Philip W.L. (2011a). "Relationship-based access control: protection model and policy language". In: CODASPY '11. ACM. DOI: 10.1145/1943513.1943539.

Fong, Philip W.L. and Ida Siahaan (2011). "Relationship-based access control policies and their policy languages". In: SACMAT '11. Innsbruck, Austria: ACM. DOI: 10.1145/1998441.1998450.

Fong, P.W.L. (2011b). "Preventing Sybil Attacks by Privilege Attenuation: A Design Principle for Social Network Systems". In: *IEEE Symposium on Security and Privacy*. DOI: 10.1109/SP.2011.16.

Gates, Carrie E. (2007). "Access Control Requirements for Web 2.0 Security and Privacy". In: *Proc. of Workshop on Web 2.0 Security & Privacy (W2SP 2007)*.

Kruk, Sebastian et al. (2006). "D-FOAF: Distributed Identity Management with Access Rights Delegation". In: ASWC 2006. DOI: 10.1007/11836025_15.

Viswanath, Bimal et al. (2010). "An analysis of social network-based Sybil defenses". In: SIGCOMM '10. ACM. DOI: 10.1145/1851182.1851226.

Yu, Haifeng et al. (2006). "SybilGuard: defending against sybil attacks via social networks". In: *SIGCOMM Comput. Commun. Rev.* 4. DOI: 10.1145/1151659.1159945.

Yu, Haifeng et al. (2008). "SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks". In: *IEEE Symposium on Security and Privacy*. DOI: 10.1109/SP.2008.13.