Identification Services for Online Social Networks (OSNs) Extended Abstract

Elena Ferrari(✉)

Department of Theoretical and Applied Sciences,
University of Insubria, Via Mazzini 5, Varese, Italy
elena.ferrari@uninsubria.it

Abstract. On-line Social Networks (OSNs) have dramatically changed how users connect, communicate, share content, and exchange goods and services. However, despite all the benefits and the flexibility that OSNs provide, their users become more reliant on online identities with often no means to know who really is behind an online profile. Indeed, to facilitate their adoption and encourage people to join, identities in OSNs are very loose, in that not much more than an email address is required to create an account and related profile. Therefore, the problem of fake accounts and identity related attacks in OSNs has attracted considerable interest from the research community, and resulted in several proposals that mainly aim at detecting malicious nodes that follow identified and formalized attack trends. Without denying the importance of formalizing Sybil attacks and suggesting solutions for their detection, in this extended abstract we also consider the issue of identity validation from a user perspective, by briefly discussing the research proposals aiming at empowering users with tools helping them to identify the validity of the online accounts they interact with.

An OSN is an online service that provides a virtual socializing platform through which people can connect to each other and share information. OSNs allow users to create personal profiles by which they present themselves to the rest of the network, share knowledge and different types of information, and connect with a large network of users (either friends or strangers). To facilitate their adoption, OSNs use very loose identity validation mechanisms: usually, only a valid email address is required for a new user to join an OSN. This way of managing identities has the advantage of encouraging users to join, however, it poses the problem of fake identities that can result in several threats to OSN users security and privacy. One of the most notable identity-related attack is represented by Sybils, a set of bogus accounts that represent multiple different identities, while in reality they are created, manipulated, and owned by one single entity (e.g., a user or a bot). Sybils can be created for different malicious aims, such as spreading malware, spying on users activity and personal information, or polluting the environment with fake content or fake support for a given cause, such as increasing the reputation of a politician or of a brand [6].

M. Hansen et al. (Eds.): Privacy and Identity 2017, IFIP AICT 526, pp. 240–242, 2018.
https://doi.org/10.1007/978-3-319-92925-5_16

Given their large impact, Sybil attacks have been widely studied in the literature from different perspectives [3], leading to the definition of both automated and human-assisted techniques [1]. Although these approaches use a variety of techniques to discriminate between Sybil and non Sybil node, they can be classified into three main groups: *Sybil prevention*, *Sybil detection*, and *Sybil tolerance* approaches. Sybil prevention aims at preventing Sybil creation. This can be mainly achieved by strengthening the identity validation process, a solution which is rarely adopted in practice due to the possible user resistance and privacy implications. The most crowded class of approaches is represented by Sybil detection techniques, which do not prevent Sybil to be created. Rather they try to identify Sybil accounts as soon as possible. Detection techniques can be further classified into *topology-based*, *activity-based*, or a mix between the two. Under the topology-based approaches, the topological structure of the underlying OSN graph is studied to discriminate between honest and malicious users. These approaches work under the assumption that the regions of honest and Sybil nodes are two separate parts in an OSN graph, where connections between them exist through attack links. However, some studies conducted on real OSN datasets have shown that this assumption does not always hold [6]. Therefore, activity-based approaches have also been explored, where the idea is to find behavioral features (e.g., frequency of sending friend requests, frequency of making new friends, ratio of outgoing to incoming activities) that could be used to reliably differentiate between fake and real accounts, using automated classifiers. Other proposals have developed a hybrid approach, combining both topological and behavioral features (e.g., [4,5]). Finally, Sybil tolerance approaches do not prevent Sybil creation nor they explicitly try to identify Sybils. Rather, they try to avoid that Sybil accounts can be used to perform malicious activities (e.g., by limiting spam in messaging systems).

All the above techniques provide good results when Sybils match known patterns of misbehavior, however they are vulnerable to some extent against adaptive attackers and their infiltration into honest communities (which is always the case in today OSNs). Therefore, to mitigate this risk, researchers have started to explore complementary services wrt the Sybil detection ones, with the aim to provide honest users with indications on the validity of other strangers profiles they plan to interact with. This would help them to take better decisions in establishing new friendship relationships, and, therefore, they mitigate the risk of fake account infiltration into honest communities. An example in this direction is represented by [2]. In this paper, starting from the observation that undetected Sybils by automated tools are the ones that succeed in emulating real profiles features and/or in getting enough connections within specific honest user communities [3], the authors study how information on a profile could be exploited to evaluate its trustworthiness within OSN communities. The developed system exploits community effort to collaboratively estimate the validity of OSN users' identities based on the coherence of some of the information they provide on their profiles.

However, since OSNs continue to grow, both in terms of their user base and the heaviness by which they are used, attackers are showing an increasing degree of sophistication in emulating real accounts. As such, we believe that identity management will continue to be an ever challenging problem, also because of the interplay with privacy and efficiency issues, that will require more efforts from both the research and industrial communities.

References

1. Al-Quirishi, et al.: Sybil defence techniques in online social networks: a survey. IEEE Access **5**, 1200–1219 (2017)
2. Bahri, L., Carminati, B., Ferrari, E.: COIP - continuous, operable, impartial, and privacy-aware identity validity estimation for OSN profiles. ACM Trans. Web **10**(4), 23:1–23:41 (2016)
3. Kansara, K.B., Shekokar, N.M.: At a glance of sybil detection in OSN. In: Proceedings of IEEE International Symposium on Nanoelectronic and Information Systems (2015)
4. Laleh, N., Carminati, B., Ferrari, E.: Risk assessment in social networks based on user anomalous behaviour. IEEE Trans. Dependable Secur. Comput. **15**, 295–308 (2016)
5. Li, Y., Martinez, O., Chen, X., Li, Y., Hopcroft, J.E.: In a world that counts: clustering and detecting fake social engagement at scale. In: Proceedings of the 25th International Conference on World Wide Web (2016)
6. Yang, Z., Wilson, C., Wang, X., Gao, T., Zhao, B.Y., Dai, Y.: Uncovering social network sybils in the wild. ACM Trans. Knowl. Discov. Data (TKDD) **8**(1), 2 (2014)