A Threat Model Approach to Threats and Vulnerabilities in On-line Social Networks

Carlos Laorden, Borja Sanz, Gonzalo Alvarez, Pablo G. Bringas

Abstract On-line Social Networks (OSN) have become one of the most used Internet services. However, as happens with every new technology, they are prone to several security issues. Despite privacy concerns begin to emerge, there are still other dangerous vulnerabilities that affect security and threaten organisations and users assets. In this paper, we present the first Threat Modelling approach in On-line Social Networks that intends to identify the threats and vulnerabilities that can be exploited. Next, we define what we call the *Circle of Risk* (CoR), a graphical definition of every security aspect involved in the threat modelling.

Key words: On-line Social Networks, threat modelling, privacy, web security

1 Introduction

On-line Social Networks (OSN) represent one of the most used Internet services, with a spectacular number of users growth, surpassing information gatherers like *Google*, *MSN* or *Yahoo!*, consuming most of the time that users spend connected to the Internet. Because there is no accepted and universal definition for OSN, this paper refers to the modern OSN that *INTECO*[1] and the *Agencia Española de Protección de Datos*[2] define in their *Study on the Privacy of Personal Data and on the*

Carlos Laorden, Borja Sanz, Pablo G. Bringas
Laboratory for Smartness, Semantics and Security (S^{3}Lab), University of Deusto, Bilbao, Spain
e-mail: {claorden,borja.sanz,pablo.garcia.bringas}@deusto.es

Gonzalo Alvarez
Instituto de Física Aplicada, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain, e-mail: gonzalo@iec.csic.es

[1] http://www.inteco.es

[2] http://www.agpd.es

Security of Information in Social Networks[3] as: 'services that let their users to create a public profile where they can introduce personal data and information. The users have different tools to interact with each other.'

Therefore, the main features of a Social Network and their tools are the *popular three C's*: Communication (allow sharing knowledge), Community (help finding and integrating communities), and Cooperation (provide tools to develop activities together).

Unfortunately, along with the aforementioned benefits come several threats. Some risks, such as social engineering techniques [11], are even exacerbated due to the excessive trust given to messages coming from friends, contacts or followed people within the OSN. In fact, OSN are one of the main significant channels to identity theft and information leaking [16, 6, 2, 5]. Furthermore, spam sending and *malware* distribution through Social Networks are increasing at an incredible pace [9, 7]. However, they are not the only threats.

The growth of the OSN phenomenon can not be ignored, neither can be integrated into the business model without knowing the risks. Notwithstanding, this expansion has transformed OSN into important applications within the world wide web, becoming the favourite target for cybercriminals. This attention requires an intensive focussing of web security efforts.

The remainder of this paper presents the first classification of the most significant threats along with the common vulnerabilities exploited, and is organised as follows. Section 2 describes what *Threat Modelling* (TM) is, offering an approximation to frequent terms. Section 3 presents the assets at risk by OSN. Section 4 details the threats that affect and compromise the assets through OSN. Section 5 discusses some of the vulnerabilities corresponding to the different existing threats. Finally, Section 6 concludes and outlines the avenues of future work.

2 Threat Modelling

Threat modelling is a description of a collection of security aspects, a set of plausible attacks which are able to affect the performance of any computer system. This methodology allows security experts to identify security risks, and develop countermeasures in the design, coding, and testing phases [13]. Therefore, analysing and modelling the potential threats that an application faces is an important step in the process of designing a secure application [3].

Being the main objective of threat modelling to provide useful guidelines on how to mitigate the associated risks, we must be able to distinguish the elements corresponding to what we have called the *Circle of Risk* (CoR) (shown in Fig. 1). The CoR is composed of *assets*, which are compromised by *threats*; threats that exploit *vulnerabilities*, which when misused result in *exposure*, which represents a serious *risk*. Finally, the *countermeasures* mitigate the dangers caused by those

[3] http://www.inteco.es/file/vuiNP2GNuMjfCgs9ZBYoAQ

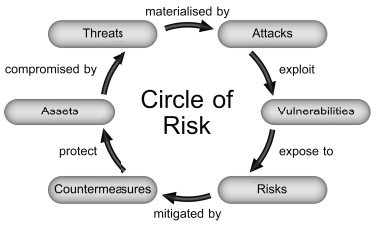

Fig. 1 Threat modelling's Circle of Risk.

risks; countermeasures which have as goal protecting the assets. Next we provide some definitions for these terms found within the technical dictionaries [4] and [12]:

- *Asset*: entity of value to the business or enterprise, be it a computer processor, disk, network link, program, datum, or user.
- *Threat*: any circumstance or event with the potential to cause harm to a system in the form of destruction, disclosure, modification of data, and/or *denial of service*.
- *Exploit*: specific attack or vulnerability used to take advantage of a particular loophole or weakness in security measures.
- *Vulnerability*: weakness in system security that could be exploited to violate system security policy; the possibility of an exploit or exposure to a threat, specific to a given platform.
- *Exposure*: proximity and/or contact with a source of a disease agent or computer virus in such a manner that effective transmission of the harmful effects of the agent/virus may occur.
- *Risk*: expectation of loss expressed as the probability that a particular threat will exploit a particular vulnerability with a particular harmful result.
- *Countermeasure*: any action, device, procedure, technique, or other measure that reduces the vulnerability of or threat to a system.
- *Attack*: the act of trying to bypass security controls on a system. The degree of success depends on the vulnerability of the system or activity and the effectiveness of existing countermeasures.

Although the threat modelling process requires the study in detail of every above-mentioned element, in this paper we introduce a first approach to the CoR, focussing on the assets, threats and vulnerabilities.

3 Assets at risk by OSN

Every enterprise has at disposal several assets that must be protected to guarantee the proper course of its business. Furthermore, following Vicente Aceituno's *Information Security Management Maturity Model* (ISM3) [1], security is defined as: 'the result of the continuous meeting or surpassing of a set of objectives'. The

loss, theft, destruction, reduction or damage of any of these assets could prevent the organisation to achieve its objectives.

Therefore, among the assets specially threatened by OSN we can identify: 1) *private information*, which can be stolen or utilised against its legitimate owner in order to harass, extort or send hypercontextual advertising; 2) *financial assets*, stolen through on-line banking fraud, telephone fraud or lost by decreased productivity; 3) *intellectual property*, which can be stolen, plagiarised or illegally distributed free of charge, causing economic losses; 4) *corporate secrets*, causing economic losses, reputation damage or decreased competitiveness if stolen; 5) *physical security*, which can be compromised by stalkers, harassers, criminals or thieves; 6) *computing and network resources*, which can be consumed leading to denial of service or decreased *Quality of Service* (QoS); 7) *corporate and personal reputation*, which can be irreversibly damaged; 8) *digital identity*, that can be spoofed or stolen.

In conclusion, the misuse of OSN affect the aforementioned assets, which are compromised by attackers who might materialize several threats.

4 Threats in OSN

Through our study, we have identified eight main categories to place every OSN threat found. Next we present our classification for threats in OSN.

- **Private Information Disclosure**. Private data management is a key feature inside Social Networks. Several threats may be included under this category:

  1. *Digital dossier building*: Everyone can collect published information through different OSN about one user and extract a complete dossier about it.
  2. *Secondary Data Collection*: Users may grant the platform secondary information (e.g., IP address, contacts lists, messages or visited pages) that without their knowledge is exploited commercially.
  3. *Reidentification*: In spite of using fake data to register on a Social Network, it is feasible to associate users shared data (i.e., text, photos, videos) with real names or e-mail directions [10].
  4. *Sensitive attribute inference*: Through machine learning algorithms it is possible to infer sensitive data from one user. Collecting data from users contacts, an experimented attacker can predict some not explicit data. For example, if all the friends of one user belong to a political party, it is safe to predict that the user has the same political tendencies.
  5. *Excessive exposition of private data*: Inadequate privacy configuration settings may lead to publishing sensitive information.
  6. *Lack of control over published private data*: Although privacy configuration settings may be adequate, it is impossible to control the information published by other users (e.g., labelled photos and videos providing unauthorised names), and even self published.

- **Financial Loss**. OSN have become a perfect channel for scams and frauds thanks to all the published personal information. Moreover, the *raison d'être* of a company is making money, and, as a money maker, the company must obtain the maximum profit of each asset. Thus, assuming that workers are one of the most important assets for the company, productivity losses caused by wasted time on Social Networks during the working day should be prevented.
- **Intellectual Property Theft**. There are two main threats concerning intellectual property. On the one hand, user's irresponsibility can lead to the publication of information protected by intellectual property rights. For example, an employee may publish a core part of an ongoing project, either because of carelessness or in order to harm the company. On the other hand, there is a lack of control over published information due the terms of use that abet the OSN, often transferring all rights to use or distribute the digital contents posted on the platform.
- **Corporate Secrets Theft**. The users' public information can provide a wide view of the enterprise where they work. Thus, an attacker could create a whole personality apparently working on the enterprise, and gain access to private data within the organisation. Furthermore, users can publish confidential information without minding the consequences.
- **Physical Security Compromise**. The over-sharing of information may lead to compromising not just digital identities but physical security. An illustrating example is *Please rob me*[4], an on-line service showing empty houses thanks to the messages posted by their owners on OSN like *Twitter* or *Foursquare*[5]. However, it is not necessary to publish explicit information or join this kind of services. *Content Based Image Retrieval* (CBIR) [15] allows inferring where a photo has been taken, hence, determining the locations that the user frequents.
  Moreover, harassment between adults, cyber-bullying (harrasment from child to child), or cybergrooming (harrasment from adult to child), educe the serious dangers connected to this threat.
- **Computer and Network Resources Consumption**. Proliferation of malware within Social Networks [8] originates a new generation of botnets that make use of the infected computer resources, such as CPU cycles or bandwidth, in order to benefit the attackers. Moreover, Web 2.0 services are based on multimedia items, whose transference through the network require a lot of bandwidth, producing productivity losses.
- **Digital and Real Life Reputation**.

  1. *Automated campaigns to erode reputation and damage image*: Attacks to undermine the reputation of the target by publishing harmful content, which can be automated due to account creating lack of control (e.g., *Sybil* attacks [14]).
  2. *Collusion*: Is an agreement between two or more users that conspire in order to undermine the reputation of a third user.
  3. *Extortion*: An attack can provide sensitive information to the attacker, information that can be used to obtain a profit through extortion and blackmail.

[4] http://pleaserobme.com/

[5] http://www.foursquare.com

4. *Repudiation*: The non-repudiation is the concept of ensuring that nobody can refute the validity of something published. This concept can not be taken into account within OSN because no one can completely guarantee the source of the published content due to profile thefts, fake profiles creation, etc.
5. *Herd effect*: Opinion leaders can polarise the judgement of thousands of users, generating a herd effect that can be used as a manipulation weapon.

- **Digital Identity**. Due to the lack of processes to verify the identity of the user when creating a new profile, fake profiles populate OSN. This problem affects especially celebrities, whose reputation and image result damaged when these fake profiles are taken as legitimate.
  Additionally, users' private data usually belong to the OSN, due to License Agreements. Therefore, OSN's negligence, or insufficient security measures may leak users' information.
  In a similar vein, identity theft attacks are used to access personal profiles and to impersonate their owners. Specifically, this fraud is commonly used to steal money or perform all kind of criminal acts. As a consequence, the victim may be pursued by the law due to the attackers' actions.

5 Vulnerabilities in OSN

This section introduces vulnerabilities commonly exploited by attackers seeking users private information.

- **Vulnerabilities associated to the platform**

1. *Difficulty to completely remove all user information when deleting an account.* When users try to leave a Social Network, license agreement clauses appear, rights that are transferred to the platform when the content is uploaded. Thus, if one would like to remove their uploaded material, it would find that the only way to do so is by deleting the videos or photos one by one manually.
   However, photographs or videos in which users are tagged do not belong to them, so the only solution is reporting the contents as inappropriate, and wait for the owner or the OSN to remove the material.
2. *Weak authentication method.* Authentication methods on the Internet are one of the most important vulnerabilities that web environments have nowadays. The combination of user-name and password is commonly misused by the user who seeks easy-to-remember login details (i.e., short user-name and passwords, passwords with no combination of numbers and letters, same user-name and password for several domains, etc.).
3. *Non validation of users data during registration process.* Most of the OSN do not use a validation process during new users registration. Unfortunately, just checking a valid e-mail address, the preferred validation requirement, is not an

adequate method, which leads to the proliferation of fake profiles populating the network.

- **Vulnerabilities associated with the data**

  1. *Disclosure of navigation data.* Communication protocols provide lots of information that users, unknowingly, send to OSN. This information provides details about users' operating systems, browsers, IP addresses, etc.; information that can be used by attackers to take advantage of the vulnerabilities that can be exploited in the victim's computer.
  2. *Information disclosed by the user status. Instant Messaging* (IM) programs and many other OSN applications provide information about users whereabouts. For example, if the user status is off-line when in that period of time it usually is on-line, the attacker knows that something unusual is going on. This also leads, to providing attackers an easy way to exploit previously found vulnerabilities when the user is away from the computer.

- **Vulnerabilities associated with the photographs**

  1. *Tagging by others.* One of the most useful features on social networks is tagging. Unfortunately, this feature also provides an easy way to find all the photographs in which one user appears, including the embarrassing or inappropriate ones.
  2. *Implicit information within multimedia content.* Most of the OSN allow the uploading of multimedia material. Users make frequent use of this feature, but they are unaware that the content uploaded contains additional meta-data. This meta-data provides details such as the camera with which the photograph was taken, where it was taken (through GPS coordinates) or when it was taken. Nevertheless, if the user removes the meta-data, several algorithms allow discovering, based on recognisable elements in the picture, the place where the photo was taken.
     Furthermore, facial recognition systems allow identifying a person on a large amount of photographs. These algorithms, combined with other technologies, allow finding singular persons on OSN with an acceptable accuracy.

6 Conclusion

On-line Social Networks represent one of the last and most important Internet services. Albeit most of the enterprises hesitate to integrate OSN into their business model, this new phenomenon can not be ignored, but neither can be adopted without knowing the risks.

In this paper, we presented a first approximation to an OSN Threat Modelling that discovers the first elements to take into account when attempting to protect a system. To that end, we identify the assets at risk, the threats that compromise them, and we note the vulnerabilities exploited by those threats.

The future work of this OSN TM is oriented in three main directions. First, we will complete the aforementioned 'Circle of Risk (see figure 1), with the attacks that materialize the threats, the risks to which assets are exposed to and the countermeasures that mitigate their effects. Second, we plan on developing a taxonomy which organises all the existing OSN threats, attacks, vulnerabilities and countermeasures. Finally, we will study the feasibility of adding weighted variables to the taxonomy in order to help identifying assets at risk to support the hardening of a system.

References

1. Aceituno, V.: ISM3: Information security management maturity model - handbook. Tech. rep. (2007)
2. Chen, B., Kifer, D., LeFevre, K., Machanavajjhala, A.: Privacy-Preserving Data Publishing. Foundations and Trends in Databases **2**(1-2), 1–167 (2009)
3. Desmet, L., Jacobs, B., Piessens, F., Joosen, W.: Threat modelling for web services based web applications. In: Eighth IFIP TC-6 TC-11 Conference on Communications and Multimedia Security (CMS 2004), pp. 161–174. Springer (2004)
4. Gattiker, U.E.: The Information Security Dictionary: Defining The Terms That Define Security For E-business, Internet, Information And Wireless Technology (KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE). Kluwer Academic Publishers, Norwell, MA, USA (2004)
5. Gómez Hidalgo, J.M., Martín Abreu, J.M., Nieves, J., Santos, I., Brezo, F., Bringas, P.G.: Data leak prevention through named entity recognition. In: In Proceedings of the 1st International Workshop on Privacy Aspects of Social Web and Cloud Computing (PASWeb) (2010). In press
6. Krishnamurthy, B., Wills, C.: On the leakage of personally identifiable information via online social networks. In: Proceedings of the 2nd ACM workshop on Online social networks, pp. 7–12. ACM (2009)
7. Luo, W., Liu, J., Liu, J., Fan, C.: An analysis of security in social networks. Dependable, Autonomic and Secure Computing, IEEE International Symposium on **0**, 648–651 (2009). DOI http://doi.ieeecomputersociety.org/10.1109/DASC.2009.100
8. Mansfield-Devine, S.: Anti-social networking: exploiting the trusting environment of Web 2.0. Network Security **2008**(11), 4–7 (2008)
9. Mazur, Z., Mazur, H., Mendyk-Krajewska, T.: Security of Internet Transactions. Internet-Technical Development and Applications p. 243 (2009)
10. Phillips, P.: Support vector machines applied to face recognition. Advances in Neural Information Processing Systems pp. 803–809 (1999)
11. Scheeres, J., Mills, R., Grimaila, M.: Establishing the Human Firewall: Improving Resistance to Social Engineering Attacks. In: The 3rd International Conference on Information Warfare and Security: Peter Kiewit Institute, University of Nabraska, Omaha USA: 24-25 April 2008, p. 325. Academic Pub. (2008)
12. Slade, R.: Dictionary of Information Security. Syngress Media Inc (2006)
13. Swiderski, F., Snyder, W.: Threat modeling. Microsoft Press Redmond, WA, USA (2004)
14. Yu, H., Gibbons, P., Kaminsky, M., Xiao, F.: Sybillimit: A near-optimal social network defense against sybil attacks. In: IEEE Symposium on Security and Privacy, pp. 3–17. Citeseer (2008)
15. Zhang, M.: Content-based Image retrieval. Artificial Intelligence for Maximizing Content Based Image Retrieval p. 115 (2009)
16. Zheleva, E., Getoor, L.: To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In: Proceedings of the 18th international conference on World wide web, pp. 531–540. ACM New York, NY, USA (2009)