CSC2012 Honours project

Threat Models in Social Networks

Ratshidaho Rotondwa Wayne

RTSROT001

Table of Contents

Abstract

Social Networks have become the most popular Internet service around the globe. Social networks are being integrated into companies and organisations to help people stay in touch, help in group or cooperative task and develop marketing and public relations campaigns[1]. Although these social networks have many benefits and advantages, they are prone to several security vulnerabilities, with privacy being the major issue. ISO 27005[2] defines vulnerability as a weakness of an asset that can be exploited by one or more threats. This literature synthesis presents Threat Modelling in Social Networks, more specifically focusing on identifying attacks against Social Network user's and possible solutions to this security issues.

1. Introduction

Social Networking web applications are amongst the most visited sites on the Internet [1], with most people spending most of their times online on this Social Networks[3][1][4]. Examples of Social Networks includes Facebook (+/- 800 million registered users)[4],Twitter, MySpace, Khoro.com to name but a few. There is no accepted universal definition for Social Networks[1], but in this paper we shall use a definition provided by Boyd, Danah m and Ellison, Nicole B [5], which provide the definition of a Social Network site:

A Social Network sites is a web-based service that allows 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
3. View and traverse their list of connections and those made by others within the system.

Therefore we can summarise the main features of a Social Network as the popular three C's[1],

1. Communication – Allow sharing massage.
2. Community – Provide tools to develop activities together.
3. Cooperation – Provides tools to develop activities together.

These social networks also gather user's personal information, which has led to an increased interest from cybercriminals. In 2009, Social Networking was the main cause of identity theft and information leaking[1].

The next section introduces Threat modelling, which is a method of evaluating the vulnerabilities in a system in order to design or implement countermeasures to protect the system against unauthorised access to information.

2. Threat Model

Threat modelling identifies formal and informal entry points, privilege boundary definition as well as the threat visualizations[6]. Entry points are places/access points where data enters or exits the application, be it authorised or unauthorised. Privilege boundary mapping is the assignment of access rights to system objects[6].A threat visualization is a representation of system attributes. A good threat visualization must capture both system specific attributes and attacker time specific details[6]. The threat modelling process is conducted during the application design phase, and it is used to identify reasons and methods an adversary would use to identify vulnerabilities or threats in a system[7], as well as developing countermeasures. According to Steven F Burns [7], a Thread model accomplishes the following tasks:

- Defines the security of an application
- Identifies and investigate potential threats and vulnerabilities
- Brings justification for security features at both the hardware and software levels for identified threats.
- Identifies a logical thought process in defining the security of a system
- Results in finding architecture bugs earlier and more often.

Approaches to threat modelling include:

1. Attack-Centric(AC)
2. Systems-Centric or Threat-Centric(TC)

Attack-Centric approach

The attack centric (AC) approach focuses on the identification of all possible entry points to the system, and possible adversary aims, in other words, we try to view the system as if we are the attacker. The attacker's actions can be categorised into one or more of the following categories[6]:

- Spoofing – The attacker pretends to be someone who has certain privileges in a system.
- Tampering – Modifies data within the system to achieve a malicious goal.
  In April 2012, it was revealed that some hackers were able to reset the Microsoft Hotmail's account password for several users. This gave the hackers access to users' email account and locked the real email account owner out. The attack was achieved by using a Firefox add-on called Tamper Data which simply views and modifies HTTP/HTTPS headers and posts parameters[8].
- Repudiation – Repudiation threats are associated with users who deny performing an action without other parties having any way to prove otherwise; for example, a user performs an illegal operation in a system that lacks the ability to trace the prohibited operations.
- Information Disclosure – Exposing protected data to a user who does not have the authorisation to view the data.
- Denial of Service – Make the service unavailable to the intended users.

Most AC-based models are mainly visualized as attack trees, hence they are simple to understand and interpret[6].

Attack Trees

Attack tree describes a directed graph which presents the why and how the security of a system can be compromised[6]. In an attack tree, every node represents an adversary goal and the root node represents the overall goal. Intermediate nodes in the graph represent sub-goals which an adversary has to accomplish in order to achieve the main objective. Leaf nodes represent goals or sub-goals that cannot be refined any further, referred to as the atom of an attack. Attack trees have simple semantics to allow the propagation of costs an adversary must incur to achieve a given task, however, semantics for attack trees have limited internal structure about and cannot facilitate sufficient logical reasoning about the threats they represent[6]. Example, when are two attack paths equal? Which event can have more impact? Attack trees have an advantage of simplicity of presentation, hence easy to interpret.

Threat-Centric approach

Threat-Centric(TC)focuses on capturing system design and deployment flaws which can translate into vulnerabilities[6]. TC approach provides mechanism of examining system design principles and deployment configuration. The threat analyst must step through the system design and deployment looking for vulnerabilities against each component of the design. Unlike AC-based model which have some semantics, most TC-based threat visualizations lack adequate semantics to allow reasoning about threats and their eventual validation. Thus, in order to use the TC-based model, the threat analyst must have sufficient background information about the system. Most TC-based models are visualized as fault trees of the system.

Fault Trees

Fault trees are a graphical representation of interaction of system failures. The failures represent system vulnerabilities which present threats to the system. A node in a fault tree represents an event and edges represent a casual-effect relationship between events. None leaf nodes represent identified hazards for which availability of data is required. Intermediate nodes leaf nodes represent refinements of a given fault. On top of Fault trees lacking expressiveness, fault tree also lack adequate semantics to facilitate reasoning about threat models. The lack of expressiveness is due to their inability to capture atomic details about the threat, like attacker knowledge, goals, tools and motivation[6].

Any threat model based only on either AC or TC is flawed because it is based on incomplete knowledge[6], hence for an effective threat model with all the knowledge, both AC and TC must be integrated.

Since the main objective of a threat model is to provide guidelines on how to mitigate the discovered associated risks, it is possible to distinguish elements corresponding components of what is called the *Circle of Risk (CoR)* (show in Figure 1) [1].

Figure 1: Threat Modelling Circle of Risk

The CoR has assets which are composed by assets, which are compromised by threats; threats that exploit vulnerabilities, which when misused, result in exposure, which represent serious risk. And finally, the countermeasures mitigate dangers caused by those risks. The goal of the countermeasures is to protect the assets, which are discussed in the next section discussed.

3. Assets at risk in a Social Network

Every person or organisation has several assets or data that must be protected. The loss, theft, destruction, reduction, or damage of any of these data could prevent the organisation from achieving its objectives or results in privacy violations. What follows is a list of data that is at risk in a Social Network[1]:

Private information

Can be stolen or used against its legitimate owner in order to harass, extort, or send hyper-contextual advertising.

Financial assets

Can be stolen through on-line banking fraud, telephone fraud, or lost by decreased productivity. The attacker can also apply the spoofing technique as discussed in section 2 get access to a user's financial assets.

Intellectual property

It can be stolen, plagiarised or illegally distributed free of charge, causing economic losses for the organisation.

Corporate secrets:

Their leakage or theft can cause economic losses, reputation damage, or decreased competiveness.

Physical security:

It can be compromised by stalkers, harassers, criminals or thieves.

Computing and network resources:

Computer resources (Memory/ bandwidth) can be consumed leading to denial of service or decreased Quality of Service (QoS).

Corporate and personal reputation:

The reputation of an organisation or people can be damaged. For example, if a hacker uses someone's compromised account to send out disgusting messages or posts, also known as frappe, people will point fingers at the owner of the account.

Digital identity:

Someone losing his\her account to the hacker gives the hacker the account's owner identity.

After having discussed data that is at risk in social networks, one can ask a question; "*How does attackers get access to this data?*" The answer to this question is discussed in the following section.

4. Attacks in Social Networks

Social Networks allows user to upload their content, as well as having open APIs and web pages heavily loaded with JavaScript and embedded media of all descriptions. This set of environment lacks security standards and practices. As mentioned in the previous section, this section focuses on attacks aimed at the assets discussed in the section 3, this attacks will be grouped in categories corresponding to the objective they are oriented to as discussed in [1].

Private Information

Sensitive data retrieval

Attackers can collect user's personal data due to their negligence when taking their private information to the public. As an example, we can take scandal such as the one that occurred in 2010, where Ron Bowles, as security consultant wrote a piece of code to collect personal data off Facebook and publish it on one of the popular sharing site, Pirate Bay[9].

Sensitive attribute inference models

Users connected in a social network usually have related attributes. Zheleva et al.[10] introduced different attacks to infer the hidden sensitive values:

- **Friend-Aggregate model (AGG):** Looks at the sensitive attribute distribution amongst the friends of the person under question.
- **Collective classification model(CC):** Aims at learning and inferring class labels of linked objects together
- **Flat-link model (LINK):** Deals with links by flattening the data by considering the adjacency matrix of the graph.
- **Block modelling attack (Block):** Uses the basic idea behind stochastic block modelling; users from natural clusters or blocks, and their interactions can be explained by the blocks they belong to.
- **Groupmate-link model (CLIQUE**): They assume that each group is a clique of friends, thus creating a friendship link between users who belong to at least one group together.
- **Group-based classification model (GROUP):** Considers each group each group as a feature in a classifier, inferring sensitive information according to the groups a user belongs to.
- **BASIC:** Use the basis for predicting the sensitive attributes of the private profiles.

Data Mining for demographic information

Using data mining techniques to retrieve public demographic data, one could infer unpublished personal data about other users.

Automated User Profiling

Retrieval of users' sensitive data by querying social networks for registered e-mail addresses and crawling every profile found to collect personal information.

De-anonymise User

Exploiting group membership information that is available on social network sites.

Social Network Mash-ups

Links data between independently provided web services to obtain previously unforeseen inference.

Social Network Aggregator

The process of collecting content from multiple social network services[3].

Financial Assets

Cross-Site Scripting(XSS)

Occurs when an application takes untrusted data and sends it to a web browser without proper validation and escaping[11].

Cross-Site Request Forgery (CSRF)

Forces a logged-on victim's browser to send a forged HTTP request with the authentication information.[11]

Bank-customer oriented malware

Malware creates target bank customers credentials. The spread of these malwares has increased due to the use of social network as a distribution channel. E.g Koobface

Intellectual Property

Contents publication property of third parties

This is when someone publishes the contents while not being the legitimate holder of the intellectual property rights of the content.

Search engines indexation of protected contents

A search engine is a database system designed to index and categorizes internet addresses, otherwise known as URLs (for example, http://www.google.com).

Loss of control over contents when users unsubscribe from the social network

When a user deactivate his/her social network account, contents that was in her account still exists, and now the user has no control over them.

Corporate Secrets

Social Engineering

Manipulating people into giving away their confidential information using information found in social network profiles.

Spear Phishing

Is an e-mail spoofing fraud attempt that targets a specific organization, seeking unauthorized access to confidential data.

Physical Security

- **Location Inferring**
  - From recognisable place in the image/ locations shown on profile
  - IP connection. [12]
- **Facial Recognition**
- **Harassment between Adults**
- **Cyber-bullying**
- **Cyber-grooming:** Paedophile

Computing and Network Resources

- **Span and Hyper-contextualised advertising**
- **Botnets:** Attacks designed solely to disable infrastructure to those that also target people and organisation

Corporate and Personal Reputation

- **Sybil Attacks:** Creating fake identities

Digital Identity

Credentials theft

When someone loses his/her username and password to an attacker.

Profile cloning

The attacker creates a false profile to gain trust from other users.

Cross-site profile cloning

Attackers' steals user's identity from a social network they are registered to and create another profile on another social network the user is not registered to and use his identity.

The next section takes care of the major counter measures in social network. Some of the attacks described above don't have a proper countermeasure yet.

5. Countermeasures on Social Networks

This section will only discuss major countermeasures in Social networks. Countermeasures reduces the vulnerabilities in a system[1]. These countermeasures will be grouped into two main categories as in [1], *platform* and *user* countermeasures. The former countermeasure prevents attacks directed to both platforms and user, while the later is for user privacy enhancement and control.

Platform Countermeasures

1. ***Technological Security of the Platform***

   Deployment of Security/Secure Socket Layer to ensure private transmissions of data.

2. ***User's Data***

   The social Network must give users complete control over their published information.

   - User should know the intended use by the social network of both personal and published data.
   - Users should be able to apply the rights to access, rectify, cancel, and oppose to data concerning to personal preferences
   - Users should be able to prevent the publication of unauthorised data. The use of tagging mechanisms request user approval is one of the approaches aimed at the achievement of this goal.

   The Social Network must also protect data against indexation of search engines by using appropriate codification.

3. ***Author's Royalties***

   Authors' rights must be protected. The social network must provide tools that allow reporting the existence of contents protected by author's rights.

4. ***User Awareness***

   Social Network must encourage their users to know what a social network does with their personal data for.

User countermeasures

1. ***User's Behaviour***

   The user must read the terms of use and privacy Policies of the social network, before the registering process and every time any change occurs. After registering, the user must define his/her privacy settings.

2. ***Technological Concerns***

   Users should use different username and passwords to access different social networks. They should also use strong password to prevent brute force attacks, and also use updated software and operating systems.

3. ***Special Consideration for Children:***

   Parents or guardians should be consulted for every sensitive action since young users are very vulnerable.

5. Conclusion

Social Networks have become one of the most important Internet applications in people lives. In this literature synthesis, we presented a first approach to Social Networks threat modelling that highlights some of the key issues to take into account when attempting to protect a system. We also

identified the assets that are at risk, the attacks that can compromise the assets, as well as some of the countermeasures required to circumvent the attacks. We also gave a brief introduction to different types of threat modelling approaches.

References

[1] B. Sanz, C. Laorden, G. Alvarez, and P. G. Bringas, “A Threat Model Approach to Attacks and Countermeasures in On-line Social Networks.”

[2] I. Standard, “STANDARD,” vol. 2008, 2008.

[3] K. Lerman and M. Rey, “Social Information Processing in Social News Aggregation Anatomy of Digg,” no. 2, pp. 1-17, 2007.

[4] H. Jones and H. Soltren, “Facebook : Threats to Privacy,” pp. 1-76, 2005.

[5] N. B. Boyd, Danah m;Ellison, “Social Network Sites: Definition, History, and Scholarship.”

[6] D. P. Mirembe and M. Muyeba, “Threat Modeling Revisited: Improving Expressiveness of Attack,” *2008 Second UKSIM European Symposium on Computer Modeling and Simulation*, pp. 93-98, Sep. 2008.

[7] S. F. Burns, “InfoSec Reading Room Threat Modeling : A Process To Ensure Application Security,” 2005.

[8] C. Graham, “Microsoft rushes out fix after hackers reset passwords to hack Hotmail accounts,” 2012. [Online]. Available: http://nakedsecurity.sophos.com/2012/04/27/microsoft-rushes-out-fix-after-hackers-change-passwords-to-hack-hotmail-accounts/.

[9] M. Elinor, “Searchable Facebook user data posted to Pirate Bay,” 2010. [Online]. Available: (http://news.cnet.com/8301-27080_3-20012115-245.html).

[10] E. Zheleva and L. Getoor, “To Join or Not to Join : The Illusion of Privacy in Social Networks with Mixed Public and Private User Profiles,” pp. 531-540, 2009.

[11] T. O. W. A. S. Project(OWASP), “The Ten Most Critocal Web Application Security Risk.”

[12] V. Lisa, “Skype IP address security flaw,” 2012. [Online]. Available: http://nakedsecurity.sophos.com/2012/05/03/skype-security-flaw/.