Chapter XVI
Network-Based Passive Information Gathering

Romuald Thion
University of Lyon, France

ABSTRACT

The information gathering process in cyber-warfare is as important as in real warfare. Once blackhats or cyber-terrorists aimed at an organization, they need to known as much as possible about its structure, its network organization, the people working in it, their addresses, hardware and software in use: the very first step of a cyber-battleplan is to know as much as possible about the battleground and the enemy. Though social engineering is a widely spread effective technique used for this purpose, other network-based techniques can be used in order to gather as much information as possible: from DNS query to infer network topology, NSLookUp to retrieve names and e-mails to intrusive techniques such as scanning tools. All this information correlated can produce very accurate results. Nowadays, the forthcoming Google Hacking is a new extremely powerful method to retrieve sensitive information anonymously. We present basic types of non-intrusive information retrieving tools, dedicated either to web server, software or hardware digging. We also describe interesting use of the Google Search engine involving advanced queries/techniques. A set of best individual and general practices are described in order to reduce the information disclosure risks.

INTRODUCTION

The rise of the Internet has been a blessing for computer science and the world of economy. It has redefined the word "information"; the Internet is the tip of the information revolution iceberg. The information revolution implies the rise of a mode of warfare in which neither mass nor mobility will decide outcomes; it is the new concept of "cyber war." It means trying to know everything about an adversary via network interconnections, while keeping the adversary from knowing much about him or herself. This tactical principle has already been exposed by Tzu in his *Art of War* (1910), but clearly it takes a new dimension in our interconnected world:

... what enables the wise sovereign and the good general to strike and conquer, and achieve things beyond the reach of ordinary men, is foreknowledge. Now this foreknowledge cannot be elicited from spirits; it cannot be obtained inductively from experience, nor by any deductive calculation. Knowledge of the enemy's dispositions can only be obtained from other men. (Chapter XIII, verses 4, 5 and 6)

In this topic, we are specifically interested in *passive* network-based information gathering techniques. In the context of networks, passive refers to techniques that do not connect to the targeted system or that would not be normally associated to an attack, whereas *active* refers to techniques that create network traffic and could be associated with suspicious or malicious behavior (e.g., port scanning).

BACKGROUND

Penetration testers, ethical hackers, and cyber criminals conduct cyber attacks in the same way. Whereas penetration testers are reliable people paid by an organization to conduct a security audit by "attacking" the target to find vulnerabilities and security weaknesses, cyber criminals and "nonethical" hackers conduct attacks without an organization's consent, to earn money, to undermine the credibility of the target, or for any other motive. In both cases, the techniques are identical. An attack can be roughly separated into five steps (FX et al., 2004).

1. **Information gathering:** By gathering as much information as possible about the target, in this step, the hacker is looking for potential vulnerabilities as well as software and hardware in use, network topology, and any information that will be useful for its attack (Grubb, 2004).
2. **Exploitation:** Using foreknowledge, a cyber criminal can focus on a specific vulnerability to take the initiative. In this step, the hacker is trying to find the most powerful and least difficult way to exploit vulnerability.
3. **Privileges elevation:** Often an exploited vulnerability does not award full control of the system. In this step, the hacker elevates his privileges to root around any means available.
4. **Cover tracks:** Once a system has been compromised, the hacker wants to cover his tracks as soon as possible, thus providing more time to act and lessen the possibility of being caught.
5. **Carry out his objective:** The hacker reaps the fruits of his or her efforts. He or she can gather any sensitive information wanted, use the compromised system to attack another one, delete data, and so forth. The hacker achieves the attack objectives.

This topic focuses on the very first step of any attack, information gathering, also known as preassessment information gathering. During this phase, the attacker is interested in obtaining preliminary information about the target—the foreknowledge.

Information gathering techniques can be roughly classified into the following:

- **Social engineering:** These nonnetwork-based techniques are the practice of obtaining confidential information by manipulating users. A social engineer fools people (e.g., by phone, by e-mail) into revealing sensitive information or getting them to do something that is against typical policies, using their gullibility. Social engineering is made possible because "the weakest link of the security chain is the human factor." For instance, the famous hacker Kevin Mitnick has extensively used these techniques (Mitnick, Simon, & Wozniak, 2002).
- **Active:** This includes intrusive reconnaissance that sends (specially crafted) packets to the targeted system, for example, port-scanning. Advanced network enumeration techniques avoid direct communication with the targeted host (e.g., Nmap (Fyodor, 2006)).
- **Passive:** This includes reconnaissance that either does not communicate directly to the targeted system or that uses commonly available public