Chapter 12
Network Forensics:
A Practical Introduction

Michael I. Cohen
Australian Federal Police College, Australia

ABSTRACT

Network Forensics is a powerful sub-discipline of digital forensics. This chapter examines innovations in forensic network acquisition, and in particular in attribution of network sources behind network address translated gateways. A novel algorithm for automatically attributing traffic to different sources is presented and then demonstrated. Finally we discuss some innovations in decoding of forensic network captures. We illustrate how web mail can be extracted and rendered and in particular give the example of Gmail as a modern AJAX based webmail provider of forensic significance.

INTRODUCTION

The main goal in forensic analysis is to reconstruct past events. We do this by analyzing evidence we have obtained and making inferences to deduce what occurred.

Depending on the type of investigation we may be interested in different events and might utilize different evidence sources. Usually, however, we are interested in high level information about different entities and their interactions. For example, we might be interested in emails sent from a certain user, the web sites visited or the chat messages they received.

Network traffic is an excellent form of evidence for forensics investigations, since it is the primary means for performing these high level interactions [Casey 2004]. For example, search terms and web browsing patterns are generally good indicators of intentions and knowledge of an individual.

DOI: 10.4018/978-1-60566-836-9.ch012

Often network forensics is important in the early parts of an investigation where disk forensics can not be obtained. This can be done passively without the need to alert the suspect and can be an important part of developing further investigative scope.

Clearly the ability to monitor network communications is a very powerful tool, and its use is heavily regulated by legal constraints. This work will not delve into the legal conditions of lawful interception [Commonwealth of Australia 1979]. Readers should seek legal advice in this regard.

This chapter follows a hypothetical investigation from the capture to the investigation stage. Our investigation is focused on an individual working within a remote office.

The first step in any forensic analysis is the acquisition of evidence. We examine some common techniques for acquiring network captures. We then discuss some architectural considerations regarding the point of capture in a network.

In our scenario we are not able to perform the network capture in the most ideal network location. We need to resort to traffic acquisition on the remote office's main feed. A problem for network forensics, is the use of network address translation (NAT) within the networks of interest. NAT makes it difficult to attribute the observed traffic to specific machines because all the traffic appears to originate from the same IP address.

We review novel techniques for source attribution and demonstrate how they can be used to classify the traffic into sources. We follow through with these techniques in order to identify the traffic generated by the individual under investigation among the rest of the traffic from the small remote office.

Once the traffic is reliably separated into sources it can be analyzed. We will demonstrate a number of tools which may be used in the analysis of the traffic. We cover some of the forensically relevant network protocols and discuss how they can be dissected into evidentiary information.

In particular we examine the HTTP protocol, HTML documents and how to render them. We also examine more complex web applications, such as modern webmail portals. We specifically examine the suspects Gmail emails to demonstrate some of the challenges encountered in the analysis of modern web applications.

This chapter contains a number of short scripts used to illustrate the points made. The scripts are there to encourage readers to try the analysis on their own captures in order to gain a feel of the concepts. I have chosen to use python for these scripts for clarity (Python is very readable even to those readers who are not familiar with it)[Python Software Foundation 2008]. I am using scapy - an excellent python library for dissecting and injecting network traffic [Biondi 2003]. Scapy is ideal for illustrations and makes the scripts easy to understand but since it is written in pure python it is too slow to run on captures of serious size. I am also using matplotlib as a plotting engine for visualizing the results [Hunter et al. 2008].

FORENSIC EVIDENCE ACQUISITION

Network forensics as a field bears many similarities to traditional Network Intrusion Detection Systems (NIDS). In many ways NIDS and network forensics systems appear very similar - they both collect and analyse network traffic. However, typically NIDS are deployed with different goals in mind.

A NIDS is designed to detect intrusions, or breaches of the security policy. On the other hand network forensics is typically interested in traffic which on the face of it looks normal, and complies with the security policy. For example, emails or web browsing activity may be of interest to the network forensics investigator, but would be classed as completely normal by the NIDS.