

Jun 10th, 1:45 PM

## Identifying Peer-to-Peer Traffic on Shared Wireless Networks


Simon Piel

*Department of Computer Science, University of San Francisco, [simonpiel.cs@gmail.com](mailto:simonpiel.cs@gmail.com)*

EJ Jung

*Department of Computer Science, University of San Francisco, [ejung@cs.usfca.edu](mailto:ejung@cs.usfca.edu)*

Follow this and additional works at: <https://commons.erau.edu/adfsl>

 Part of the [Computer Engineering Commons](#), [Computer Law Commons](#), [Electrical and Computer Engineering Commons](#), [Forensic Science and Technology Commons](#), and the [Information Security Commons](#)

---

### Scholarly Commons Citation

Piel, Simon and Jung, EJ, "Identifying Peer-to-Peer Traffic on Shared Wireless Networks" (2013). *Annual ADFSL Conference on Digital Forensics, Security and Law*. 3.  
<https://commons.erau.edu/adfsl/2013/monday/3>

This Presentation is brought to you for free and open access by the Conferences at Scholarly Commons. It has been accepted for inclusion in Annual ADFSL Conference on Digital Forensics, Security and Law by an authorized administrator of Scholarly Commons. For more information, please contact [commons@erau.edu](mailto:commons@erau.edu).

**EMBRY-RIDDLE**  
Aeronautical University™  
SCHOLARLY COMMONS

(c)ADFSL



# **IDENTIFYING PEER-TO-PEER TRAFFIC ON SHARED WIRELESS NETWORKS**

**(Briefing Paper/Presentation)**

Simon Piel ([simonpiel.cs@gmail.com](mailto:simonpiel.cs@gmail.com))

EJ Jung ([ejung@cs.usfca.edu](mailto:ejung@cs.usfca.edu))

Department of Computer Science

University of San Francisco

2130 Fulton St

San Francisco, CA 94117

Phone: (415) 422-5422

Fax: (415) 422-5800

**Keywords:** peer-to-peer, contraband download, tracing, investigation tool

## **ABSTRACT**

Tracing contraband downloads leads investigators to an IP address, and in turn Internet Service Providers (ISP) can provide a physical location using this IP address. However, most homes and offices share this IP address among many computers using wireless networks. In other words, there needs to be another investigation to find out which computer was responsible for contraband downloads. To make matters worse, these shared wireless networks often have vulnerabilities in access control such as using WEP or using weak passwords. In such cases, any computer in range, not necessarily at the given physical address, could be responsible. We use shallow packet analysis to identify which computer in the shared wireless network is participating in peer-to-peer downloads. Our approach does not require the packet content, thus does not require wiretapping warrant. We discuss characteristics of peer-to-peer traffic and show how we derive and use them. Our approach monitors the patterns in the duration, the frequency, the amount of information uploaded and downloaded, and the download speed in all connections. In particular, we monitor the traffic distribution over time for each connection and combine them based on their unencrypted header information to learn which connections are likely to stem from which application.

## **1. INTRODUCTION**

Contraband, such as child pornography and copyright infringement materials, is a prime target of investigations. Tracing contraband on the Internet is challenging already, but even when the investigators succeed in tracing, it often leads into a shared wireless network. Identifying which computer in the shared wireless network is responsible for contraband transfer is not trivial. Even when the owner of the wireless network is cooperative, wireless network may have vulnerabilities such as using a weak protocol (e.g. WEP) or weak passwords, and the owner may not be aware of unauthorized users. If the offender is well versed in computer technology, the culprit might be encrypting the transfer between his computer and the source of illegal material. Unfortunately, decrypting the traffic requires excessive time and resources to the point that monitoring the content of the network becomes infeasible. Also, it requires wiretapping warrant to collect the content of the traffic. We therefore forego the approach of deep-packet analysis that requires on the contents of the packets, and instead focus on the lower layers of the network stack, which are necessarily unencrypted due to their role of directing packet flow through the network. Specifically, we focus on the headers of the transport and network layers, which contain important flow and routing information. This allows us to measure the network traffic characteristics, then to make an educated guess on the type of

applications running on each computer in the shared network. Since peer-to-peer application usage has a high correlation to the contraband transfer, or we focus on identifying computers participating in peer-to-peer file transfer only with header information.

## 2. GOALS

The goals of this research are:

1. Peer-to-Peer Application's Network Characteristics  
Find and establish distinct characteristics of peer-to-peer file transfer based on the header information.
2. Tool Development  
Develop a software that analyzes the shallow packet information (encrypted traffic or traffic logs obtained without wiretapping warrant) and produces traffic patterns of each computer.
3. Guideline for identifying computers participating in peer-to-peer file transfer.  
Based on the results from step 1, provide guidelines in identifying computers in peer-to-peer file transfer.

## 3. PEER-TO-PEER TRAFFIC CHARACTERISTICS

Peer-to-peer networks have the advantage that they don't rely on a single source of distribution. This benefits the user, as the file's download speed is not limited by a single server's upload speed, but rather by the collective upload speed of all the users "seeding" the file in question. Furthermore, since the files are distributed over many seeds, there is no single point of failure that could be exploited by law enforcement to restrict access to a given file. These traits of peer-to-peer file sharing make it a popular choice for spreading and sharing contraband over the Internet. Our tool measures and plots these traits for visual inspection. The plots are easy to understand without deep understanding of network technology. We discuss each characteristic with example plots below.

### 3.1 Probing to multiple IP addresses

Because peer-to-peer applications have to probe many users to establish what chunks of the file each seed possess and whether or not the user can download from that particular seed, we can use this probing communication to multiple unique IP-addresses as a reliable indicator for peer-to-peer traffic. The probing and connecting to multiple seeds results in a large number of unique connections that

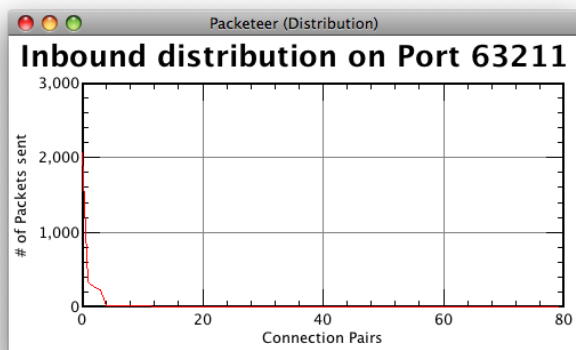


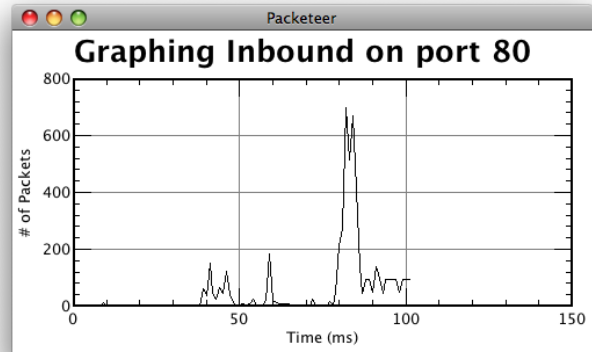
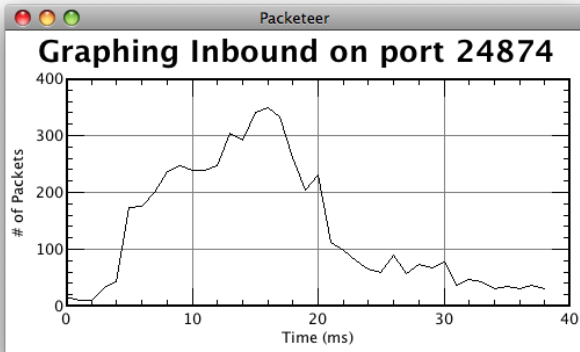
Figure 1 Long tail distribution of p2p traffic

contain a varying number of packets exchanged for each connection. Note that this probing does not only happen in the beginning of file transfer, but periodically happens to optimize transfer speed. We can visualize this distinguishing behavior by plotting the number of connections over the size of the packet exchanges. This results in long-tail distribution shown in Figure 1. X-axis shows the unique connections, and Y-axis shows the number of packets transferred in each connection. The connections are sorted by the number of packets, and show the long-tail distribution of per-connection traffic.

### 3.2 Steady Traffic over Time

When plotting the number of packets (x-axis) versus time (y-axis) for a single port and direction, we detected that peer-to-peer software typically has a slow and steady increase in packets as shown in

Figure 1, whereas other network traffic, such as video streaming or website visits, can be distinguished by sharper edges on the graph as shown in Figure 2.



### 3.3 Inbound Versus Outbound Traffic

Peer-to-peer file distribution works best with a high number of seeds, thus, peer-to-peer software encourages its users to seed files in order to increase their download speed. As a result comparing the inbound versus outbound traffic of a port will yield a similar shape on both graphs. While the height of both graphs will vary depending on the download/upload ratio the user has selected, the overall shape

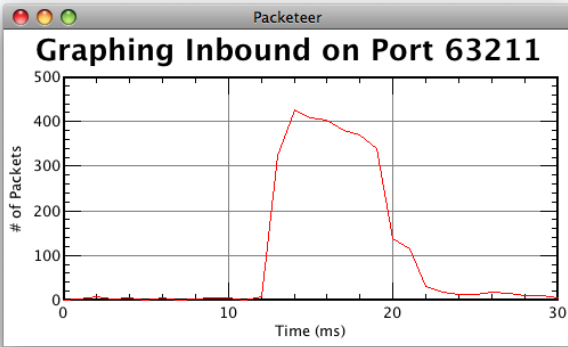


Figure 3 Inbound traffic (p2p download)

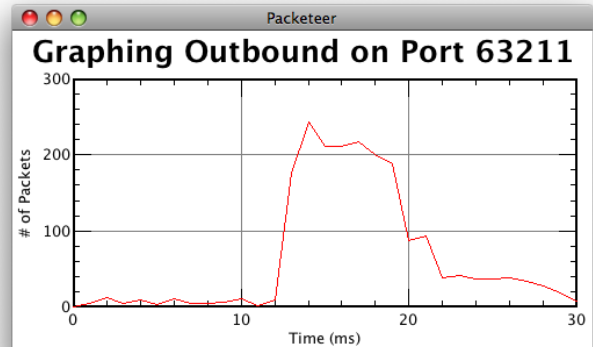


Figure 4 Outbound traffic (p2p upload)

is very similar, as shown in Figures 3 and 4.

### 3.4 More connections

We have found that a peer-to-peer application plot will have a more gradual fall compared to other network traffic. In other words, p2p traffic connects to many more IP addresses than other applications, and data transfer per connection varies. This gradual fall is shown in Figure 5.

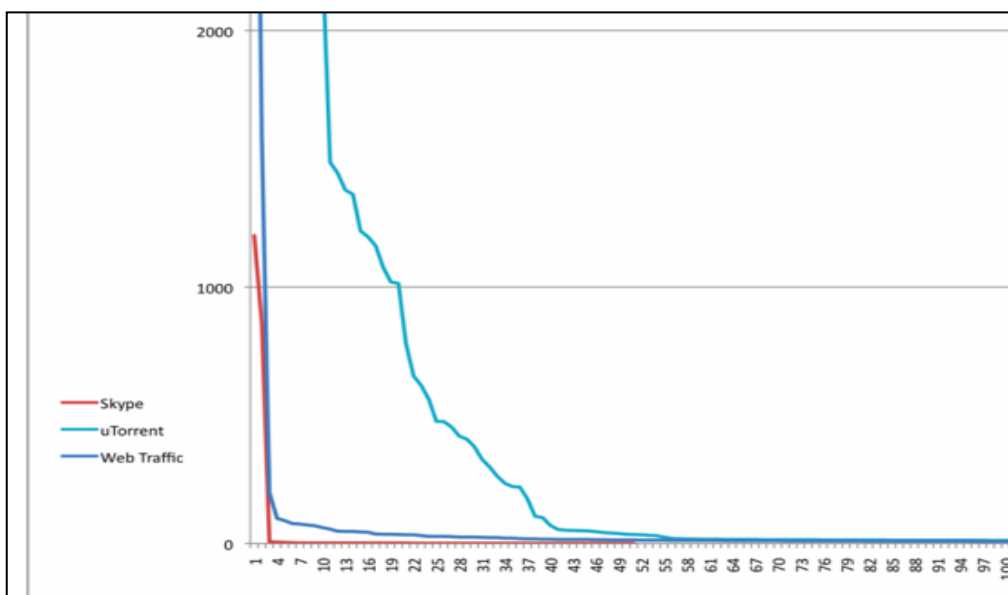


Figure 5 Gradual fall of p2p download

## 4. IMPLEMENTATION

To have full control over our capture environment, we used Wireshark to tap into and capture our own test network's traffic. We elected to build a platform-independent Java application to do the packet capture analysis and plotting. Since the number of ports getting captured is very high we choose to create an easy to maintain and easy to customize configuration file, that can be used to exclude specific ports (such as port 80 for web-traffic) or white-list ports that are commonly used by peer-to-peer software. In addition to showing the plots for the desired ports and saving them as images on disk, we also output the information in text format, so that future analysis can use these files to automate the evaluation process, or the investigator can modify the configuration file to include other suspicious ports for analysis. All our figures are produced by our own software, and show consistent characteristics of p2p file transfer.

### 4.1 Adjustment for changing ports

Most peer-to-peer software has the option of randomizing its port for each program execution. We group all traffic between a unique pair of IP addresses to detect p2p traffic on any ports.

### 4.2 Data Collection

We reorganize the data from the raw capture file (.pcap file in text format) by extracting each packet's receivers and senders IP addresses from the network layer header as well as the sender's and receiver's ports from the transport layer header. Each packet's data is then stored in data structures that can produce the plots to show the frequency and the amount of traffic of each connection.

## 5. CONCLUSION

The identification of peer-to-peer traffic through shallow packet analysis opens up a promising new way to detect computers participating contraband transfer in a shared wireless network. We have implemented a proof of concept and shown that our software provides reasonable results. The shallow packet analysis is far less intrusive than alternative deep-packet analysis methods and provides useful information even when the data transfer is encrypted. Our software is available at <http://www.cs.usfca.edu/~spiel/packeteer/>.