Lesson Plan

Lesson 1.1.4 Identifying Suspects on the Internet

Lesson 1.1.4 Identifying Suspects on the Internet	Duration: 1 hour

Resources Required:

- Laptop or PC running Windows 7, 8 or 10 and with Microsoft Office 2010 or later
- Projector and display screen
- Whiteboard, flipchart or other technique for recording student input
- These resources are only needed if the trainer is using a PowerPoint presentation

Session Aim:

To provide the delegates with information about the relationship between an individual's online activity and their real world identity, and some of the challenges associated with creating that link.

Objectives:

By the end of the lesson the students will be able to:

- Explain important Internet terminology
- Describe what is an IP Address
- Describe how can you find a suspect IP address
- Describe how can you associate a suspect IP address with a real person
- List three technical challenges to identification of suspects on the Internet

Trainer Guidance

This session provides information on the identification of suspects online and the association of an individual's online activity with their real world identity, as well as describing some of the challenges of creating that association.

All information about this session is included in the PowerPoint presentation entitled "Session 1.1.4 – Identifying Suspects on the Internet" in the resource pack. The trainer is responsible for ensuring that the materials are up to date. Changes may be made; however the objectives should be achieved.

Lesson Content

Slide Numbers	Content
Slides 2-4	The trainer should refresh the concept of an IP address. The salient points that are relevant to this lesson in particular are summarised on these slides.

Slide 6	This section of the course explains how it can be possible to associate an IP address with some suspicious activity online. This is a necessary component of an Internet based investigation.
Slide 7	Trainer should explain that web servers will typically keep a record of all incoming requests. Trainer should also explain, however, that this is a configuration option in a web server and a person who is running a web server can switch off the logging and may well do this if they are attempting to conceal the activity of the people using the web server. This will mean that no logs of the web browsing activity will be kept.
Slide 8	The trainer should explain that this is an example of a web server log. Trainer should refer to the highlighted log entry and point out: 1. The source IP address – 216.128.130.126 2. The time/date of the HTTP request – 23/Sep/2006:21:30:02 -0400 – Trainer should point out the timezone, in this case GMT-4. 3. The thing that was requested "/pgp/" – this is known as the resource and will have meaning in the context of the server that is handling the request. It might be a document on the server, a folder on the server or an image, for example. 4. The page the user was on when they clicked on the link that led to this request. This is known as the referrer. In this case the long string starting with dir.yahoo.com. From this we can tell that the user was looking at this page, clicked on a link that led to this server receiving a request for a resource named "/pgp/". The browser identification string (starting Mozilla/4.0), which is provided by the browsing software of the client PC, although this can be faked/changed.
Slide 9	Trainer should explain that unlike web browsing, the sender and recipient of an email do not interact directly.
Slide 11	Trainer should explain that SMTP (Simple Mail Transfer Protocol) is the process by which email is sent and forwarded, POP (Post Office Protocol) or alternatively IMAP is the process by which email is checked/received. Trainer should talk through the process of delivering an email via an outgoing mail server and an incoming mail server. Trainer should explain that in the case of web-based email (e.g. gmail) that instead of using POP3 to interface with the incoming mail server, that the user logs in to a web server and reads the email directly on the incoming mail server.
Slide 13	Trainer should explain that, although the students might not recognise email in this format, each email that have ever read has headers like this. Trainer should highlight in particular 1. The Received headers, which show IP addresses that processed this email 2. The from and to email addresses 3. The subject 4. The Date Trainer should point out that it is possible for some of this information to be faked if someone is trying to conceal the source of the email and therefore

	email header analysis requires some technical skill to interpret correctly.
Slide 14	Trainer should ask the students to think of other services that might have IP addresses. Trainer should stimulate discussions with examples such as Skype, Pokemon Go, DropBox, Facebook, Google, Microsoft. Etc.
Slide 15	This section of the course explains how it can be possible to associate an IP address with some suspicious activity online. This is a necessary component of an Internet based investigation.

Slide 16	Trainer should remind students that the Internet is much more complex than we can possibly hope to cover in a short introduction like this. Trainer should caution students that there is not a simple one-to-one mapping between every person in the world and an IP address on the Internet. There are many reasons why an IP address does not necessarily identify a specific individual (it is a necessary but not sufficient piece of Information).
	Two specific examples should be discussed:
	
	1. Static versus Dynamic IP addresses – Trainer should ask students to consider the case of their Internet connection at home. Most likely they either use a cable Internet provider or a DSL provider. In both cases, the trainer should explain, that the Internet Service Provider might, and most probably will, periodically change their IP address. This means that it's important to not only know what IP address was in use, but also WHEN the IP address was in use to be able to identify a particular subscriber. A similar thing can happen within organisations – each time a PC is rebooted it might get a different IP address.
	2. Network address translation – Trainer should explain that despite the fact that there are 4 billion available IP addresses, they are not allocated one by one, rather they are allocated in groups. One a group has been allocated to an organisation, no other organisation can use any IP address in that group. Trainer should explain that this leads to very inefficient allocation of IP addresses and therefore there is a problem with available IP addresses running out. Trainer should explain that this, amongst other reasons, has led to the need for a revised version of IP including longer addresses (IP version 6). Trainer should explain that there are certain groups of IP addresses that are not allowed to be used on the Internet. In other words, they will not be forwarded by the shared network infrastructure. Trainer can give the examples of any IP address starting with 10, or any IP address starting with "192.168". Trainer should explain that these IP addresses can be used perfectly well on a local network, but not on the Internet. Trainer should explain that these IP addresses are used to make more efficient use of the available IP addresses in the following way:
	    1. Within the organisation, each PC is given an IP address that cannot be used on the Internet.
	    2. The organisation has a small number, perhaps just one, real Internet IP address assigned to it.
	    3. Whenever a PC wants to communicate with another computer on the Internet, the organisation's infrastructure (Internet gateway router in particular) substitutes the real internet IP for the internal IP address before forwarding the traffic out to the Internet.
	    4. This allows a large number of PCs share a single IP address for communicating on the Internet, and the technique is known as Network Address Translation (NAT).

	The upshot of the use of this technique, however, is that all of the PCs inside the organisation appear to the rest of the Internet as if they are a single computer with a single IP address. If this IP address is identified as part of an investigation, it is only the organisation itself that is in a position to determine which internal PC generated the traffic in question.
Slide 17	Trainer should explain that one of the key sources of IP to subscriber identity information. The process for requesting such information from an Internet Service Provider will depend on national legislation, although some form of prosecutorial or court order is typically required. The period of time for which ISPs are required to retain records is also a matter for national legislation. A period of 2 years is common at the time of writing. Which ISP owns the IP address in question: https://www.whoismyisp.org/ is a helpful resource in this regard. The trainer should point out that the ISP may not be a national ISP in your country and therefore mutual legal assistance may be required to get access to this information.
Slide 18	Trainer should point out that multinational service providers typically keep records of, for example, the IP address from which a particular account logged in. Trainer should explain that in cases where you are aware of a suspect, for example, Facebook account, it may be possible to find an IP address associated with the activity of this account from Facebook. International cooperation, including cooperation with multinational ISPs is covered elsewhere in this course in more detail.
Slide 19	Trainer should describe the scenario where an investigation points to an IP address owned by a company. Trainer should point out that in many cases companies use NAT to share a single internet IP address amongst all internal computers. In these cases, only the company is in a position to associate specific IP address usage with a particular internal PC. Therefore engagement with the company to assist with the investigation will be required. How this engagement takes place will depend on national legislation. Trainer should explain that this situation can be complicated by the fact that, even though records may exist, they might not exist within your country. If, for example, a company is part of a multinational organisation, they might not control or operate their own infrastructure. In a case like this, if records exist, they may be in another jurisdiction. Whether the company in your country are in a position to provide the records will depend on the organisational and technical structure.
Slide 21	Trainer should explain that criminals may attempt to intentionally obfuscate their IP address, making it virtually impossible to identify their IP address. Trainer should also explain that certain technologies, through their normal use and operation, might also make it virtually impossible to identify the suspect IP address.
Slide 22	Trainer should explain that services exist online specifically to conceal the source of browsing and Internet activity. Trainer should show example webpage on next slide. Trainer should point out that similar services are available to anonymise the source of email.

	Trainer should explain that other services, such as TOR, the onion router, can be used to conceal source IP data.
Slide 24	Trainer should explain that Carrier grade NAT is a serious problem for identifying a suspect associated with a particular IP address. The problem is that a second piece of information, known as a port number, is required to unambiguously associate a particular user with the IP address. In many cases, the port number is not stored in service records. Trainer should refer back to the web server logs and email headers and point out that no port number information is present in either of these sources. Trainer note: See https://en.wikipedia.org/wiki/Carrier-grade_NAT for useful background information.

Practical Exercises

No practical exercises are envisaged in this lesson.

Assessment/Knowledge Check

No knowledge check or assessment is prepared for this session.