ENSec: A Theoretical Framework for Endpoint Network Security

Abderrazak Bachir Bouiadjra *
Evotionary Engineering and Distributed Information Systems Laboratory

Abstract

Executing network attacks within a LAN has become quite simple regarding to the large number of available tools that can be both downloaded and used easily. Ones of their main goals are to gain an unauthorized access to some devices, to stop some or all of their operations (denial of services) and to steal secret & confidential information using man in the middle techniques. Unfortunately, such tools are used in many networks; they exploit the rich set of vulnerabilities present in the operating principles of some well-known protocols such as: ARP and DHCP. However, the lack of standards to secure their flaws on one side, and having solutions embedded in network devices such as switches on the other side, remains a significant barrier against having the needed confidence when connecting to untrusted and/or unknown networks. In this paper, we provide a survey on dangerous network threats and attacks with existing solutions, and we propose a theoretical framework ENSec: for Endpoint Network Security in order to meet those requirements.

Keywords: Network Security, Network Threats, TCP/IP, ARP, DHCP.

1. Introduction

The easiest way to secure a network from outside attacks is to close it off completely from the outside world. In this case it will be considered as a closed network and it will provide connectivity only to interior devices. Because there is no outside connectivity, network administrators and IT managers can consider their networks safe from most of dangerous attacks. But unfortunately this is a big mistake that most of people believe in. Here are some reasons:

- Most of Security Organizations (such as CSI and NSI) estimates that 60 to 80 percent of network attacks and misuses come from inside.
- Security vulnerabilities and threats have increased significantly in last years.
- Hacking tools have become easier to find, to download and to use; in addition, to run advanced attacks they require little or no networking skills.
- Today's networks are considered as a way to earn money, which leads to ensure the maximum amount of availability, stability, confidentiality and integrity in the network.

Two major types of attacks are frequently used on networks: Denial of Service (DoS) and Spoofing Attacks.

DoS attacks focus on sending a particular traffic to a network device in order to stop some or all of its operations. Many attacks of this kind exist, some of them aim to send ICMP traffic continuously in order to slow down or stop the operation of the target. Others exploit the operating-system’s flaws and vulnerabilities in order to gain access to the target device, or at least to cause its complete-stop by sending malformed packets. Others aim to saturate the memory and/or the CPU of the target by sending a large amount of traffic with unknown source & destination MAC/IP addresses...

Spoofing attacks are also varied and aim to impersonate a device other than the legitimate one. The well-known is intended to deceive the user by forcing him to unwittingly use the attacker's machine as a default-gateway than the legitimate one, which will capture all user’s traffic to the outside of the network and redirect it to the legitimate gateway, which is known as “hidden Man In The Middle attack”. This attack aims to steal secret and confidential information such as login’s passwords and also credit-card number introduced during Internet transactions.

However, the target to achieve in the present research can be summarized in ensuring a considerable level of confidence to users when connecting to both untrusted and unknown networks, in order to prevent most of well-known network threats and attacks. Those attacks can be executed not only by some malicious users, but also by some network administrators, who have the possibility to disable some security features (during some periods) without any trace neither locally nor on accounting servers.

The rest of the paper is structured as follows. In section 2, we provide an overview of some well-known protocols and their related threats. Section 3 provides a state-of-the art literature review of current works. In section 4, we outline the main mechanisms of the theoretical framework for endpoint network security (ENSec). In Sections 5, we provide a discussion about some topics and offering answers to related questions in order to clarify fuzzy points. Finally, in section 6 we provide concluding remarks and perspectives for future enhancements.

2. Protocols and Threats

DARPA (Defense Advanced Research Projects Agency) have originally developed the TCP/IP model in order to connect networks of the United States’ Department of Defense. Afterwards, the TCP/IP was used to connect public and private institutions across the world. The TCP/IP suite includes four layers with a huge number of protocols. The most related ones to our research are:

2.1. Media Access Control Protocol

MAC protocol is used to provide the data-link layer of different technologies such as 802.3 Ethernet and 802.11 WiFi. It encapsulates payload-data of Internet layer by adding a 22 bytes header (including preamble, MAC addresses and Ethertype), 4 optional dot1q bytes and appending the all to an integrity checksum of 4 bytes. MAC addresses are 6-bytes (48 bits) in length and represented with hexadecimal format. They are divided into manufacturer ID (OUI represented in three bytes) and a serial number in the remaining three bytes, which means that they are globally unique identifiers assigned to each network interface card. They are stored in read-only memories and therefore they are often referred to physical addresses.

The most important information to know is: the MAC address is learnt only one time (during the loading step of the operating system), and it will be stored into the random access memory RAM, whence it will be used continuously during the encapsulation of network's traffic. Based on that, some available hacking tools offer the possibility to spoof the factory-assigned MAC address in order gain an unauthorized access (when MAC-filtering is enabled), and also to hide itself and to impersonate as another one. In addition, another threat is present regarding operating principles of Ethernet switches; where they populate their mac-address tables depending on the traffic that pass through them, and store an entry about the source mac address with the receiving port of each frame. As a result, attackers can run easily MAC flooding attack in order to saturate the reserved memory-space to the mac-address table with random mac addresses; at this step the Ethernet switch will act as a Hub and will broadcast all traffic to all ports, which allow the attacker to spy all traffic about the whole network's users.

2.2. Internet Protocol

TCP/IP uses the IP as the primary Internet layer's protocol for delivering traffic from a source host to a destination one solely based on the addressing methods. IP addresses are 4-bytes (32 bits) in length and represented with decimal dotted format and therefore they are often referred to logical addresses. They are divided into network and host parts using the Subnet-Mask which contains (in its binary representation) a set of ones (that represent the network-part) followed by a set of zeros (that represent the host-part).

The most important rule of IP addressing defines that hosts belonging to the same network need to have the same network-part in their IP addresses with different and unique host-part. Based on that, some available hacking tools offer the possibility to spoof the IP address of a device in order to gain an unauthorized access (when IP-filtering is enabled), to hide itself, to impersonate as another one, and also to occur an IP address conflict with other devices (such as server or surveillance camera), what render either one or both of them unusable for network operations. Such attacks combined with other techniques can result in Denial of Service.

2.3. Dynamic Host Configuration Protocol

TCP/IP uses the DHCP in order to automatically assign IP information (such as address, mask, default-gateway and DNS servers) to network devices, especially where their number is significant. DHCP eases considerably the task of assigning IP addresses; in addition, it ensures their correctness and avoids duplicate and invalid addresses.

DHCP client sends a DHCP Discover message as a broadcast, in order to find a DHCP server. DHCP server replies with a DHCP Offer message. DHCP client sends a DHCP Request in order to reserve and to use the offered address. DHCP server replies finally with DHCP Acknowledgment informing the client that the address is reserved to him during a lease duration.

This operating principle allows attackers to appear as an illegitimate DHCP server, and sends DHCP offers containing valid IP address (within the subnet) with the attacker's IP address as the default-gateway. As a result, the attacker will act as "Man in the Middle" where traffic between the victim and another host will be redirected through the attacker, without the two knowing it; allowing at the same time the attacker to spy traffic of the victim device.

2.4. Address Resolution Protocol

TCP/IP uses the ARP in order to resolve network-layer addresses into link-layer addresses (such in Ethernet: IP to MAC) [1]. When a host needs to communicate with another, knowing only its network-layer address (whether directly or from DNS), it will broadcast an ARP request to all devices on the same subnet, for getting an ARP reply containing the link-layer address of the desired device. Afterward, the sender device will be able to communicate with the receiver because it has all needed information to complete the encapsulation process (sources & destinations of link & network layers addresses). Network devices build and store ARP-Cache containing the mapping between IP & MAC addresses of all known hosts. This ARP-Cache can be refreshed using ARP announcement packets (also known as Gratuitous-ARP) that update any stored entries of any device.

This operating principle allows attackers to send those Gratuitous-ARP packets containing the attacker's MAC address as the default-gateway's MAC address. As a result, ARP poisoning "Man in the Middle" attack is done, allowing the attacker to spy traffic of the victim device. Various tools available on the Internet [2], allow attackers to perform the previous attacks.

3. Related Works

A state-of-the art literature review on solutions of previous threats and attacks is suggested in this section. We propose to classify researches and works in this area by following the order of the previous section.

Starting from MAC spoofing & flooding attacks, diverse solutions (such Port-Security) have been proposed variously by different manufacturers of network devices such as Cisco and Juniper [3,4]; but they tend to be broadly similar in term of principles: they define in their switches, a maximum number of allowed MAC address to be learned per port, and drop all frames with other addresses.

IP address spoofing has a well-known solution implemented by some manufacturers of network devices including Cisco and Juniper called: IP Source Guard [3,4]. It checks the IP source address in a packet sent from a device connected to an untrusted switch port against entries stored in database containing mapping between MAC, IP and Switch port called DHCP Snooping Database. If the switch determines that a packet contains an invalid source IP address, it drops it immediately.

Moreover, to address the DHCP spoofing attacks, several solutions have been proposed in the literature where the most important of them are: First, DHCP Snooping proposed by Cisco and also by Juniper [3,4] allowing their Ethernet switches to inspect all DHCP offers and to allow only those coming from the legitimate DHCP servers of the network (coming from trusted switch ports). In addition, during inspection of DHCP packets (discover, offer, and request) the switch populates its DHCP Snooping Database with mapping between MAC, IP and Switch port of network devices. Second, CLL: Cryptographic Link Layer [5] provides authentication and confidentiality to the hosts of the LAN using certificates which raise some cost and complexity issues to implement such solution and constitute at the same time, a new target for DoS attacks since asymmetric cryptographic operations are very expensive than symmetric ones.

However, to address the ARP poisoning attacks, several solutions have been proposed in the literature where the most important of them are: First, Static ARP methods, where every device need to have a full mapping of all others (IP & MAC). As a result, the ARP spoofing will not be allowed; but this solution is not scalable at all, it will cause a problem with mobile devices (such as laptops and smartphones); and finally, it is not efficient where operating systems (such as Microsoft Windows XP and 7) overwrite static ARP entries when receiving Gratuitous ARP replies and does not prioritize static over dynamic ARP. Second, Passive methods, where every device will monitor the ARP traffic on the network and build a complete database, and if it notices a change in any entry using Gratuitous ARP it denies the packet and report an alarm about the violation. The most popular tools in this field are ARPwatch [6] that is independent from operating system; Anticap [7] and Antidote [8], which are considered as patches of operating system's kernel. Third, Dynamic ARP Inspection (DAI) proposed by Cisco and also by Juniper [3,4] allowing their Ethernet switches to drop all ARP replies with invalid IP/MAC mapping basing on information stored in the DHCP Snooping Binding Database. Fourth, Secure-ARP [9] that uses Digital Signature Algorithm (DSA), it is considered as a replacement for the well-known ARP protocol, which leads to the necessity to upgrade the network stack of all devices. Fifth, Ticket-based-ARP [10] implements security against ARP poisoning attack by providing a centralized solution that distributes tickets containing MAC/IP addresses mapping. Sixth, Enhanced-ARP [11] which prevents this attack by rejecting all ARP refresh and announcement packets about hosts that are still alive and by using a voting-based resolution mechanism for new entries.

4. ENSec Framework

Our aim is giving some confidence to end-users by securing them against the rich set of threats and attacks described above, even if they connect to untrusted and unknown networks; whether if they use switches without security features (caused by hardware/software limitation, or disabled by the administrator). Moreover, we focus our efforts to propose a backward-compatible solution with all existing TCP/IP suite protocols, and independent of all hardware and software constraints (such as network devices and operating systems).

For those reasons, we propose five main features to meet the previous requirements: Layer2-ICMP, Layer2-TTL, Layer2-Trace, Rogue Detection and Advanced Filters.

4.1. Layer2-ICMP

The Internet Control Message Protocol (ICMP) is defined in RFC 792 as a part of the TCP/IP protocol suite. Its principle is to send-back to the source of a data-flow, some ICMP messages to monitor, control and also to respond to failures.

By the same way, ENSec solution implements a Layer2-ICMP suite basing on the same RFC and offers the possibility to use the well-known diagnostic tool Layer2-Ping using MAC addresses.

The more important features of the proposed mechanism can be summarized on its ability to work even before having Layer3 connectivity (such as: before getting an IP address from DHCP); and on its ability to discover all alive Layer2 devices using scanning and broadcasting mechanisms. In addition, rather the ICMP protocol that can be easily blocked through Firewalls we propose to use an advanced process basing on some unblocked protocols such as ARP.

4.2. Layer2-TTL

TTL (Time To Live) is an 8-bit field in the IP header, its value is set by the sender of a packet, and reduced by every router on the path, if it reaches zero the packet will be dropped and an ICMP message is sent back to the sender.

By the same way, ENSec solution implements a TTL field of 8-bits as a replacement of SFD (Start of Frame Delimiter) in Ethernet header. The proposed TTL field will play an old role as SFD using its four firsts bits and a new role as TTL using its four lasts bits. Therefore, the proposed field will contain always a binary value equal to “1011” which indicates the start of frame. In addition, the next sub-field will be a binary-value lower or equal to “1111”. As a result, the maximum number of Layer2 devices within a broadcast-domain (subnet) will be 15, which is largely sufficient in small and medium sized networks taking into account the implemented segmentation using VLANs. We propose in ENSec solution to not drop frame when their Layer2-TTL reaches zero, but rather to allow it to continue throughout the entire layer2-path; and a Layer2-ICMP message will be sent back to the sender.

The proposed mechanism can be very helpful in security field, but also in other fields as redundancy and availability, where it can be helpful to avoid network congestion caused by Layer2-Loops for example.

4.3. Layer2-Trace

The basic principle of Traceroute is sending IP packets (UDP, TCP or usually ICMP) with a TTL (Time-To-Live) starting from value “1” and will enlarged increasingly. When receiving a packet, the router decrements the TTL and does a test to verify if the TTL reaches 0 then it sends an ICMP error “TTL Exceeded” to the source. Basing on those principles, Traceroute utility discovers the nearer and nearer layer 3 devices.

Similarly, a layer 2 Traceroute Utility have been proposed in the literature by Cisco Systems [3], where they offers the possibility to discover the nearer and nearer network layer 2 devices going from a host to another. The main barrier and limitation of this solution reside on its ability to work only using Cisco devices and only if Cisco Discovery Protocol (CDP) is enabled on them.

We propose in ENSec solution to work basing on Layer2-TTL mechanism in order to discover MAC addresses of network devices on the path and offers the possibility to detect and to avoid all kinds of Man in the Middle Attacks by tracing and monitoring all devices on a path: before, during and after the communication process. When a suspect comportment is detected, we propose to do two main actions: Rogue Detection and Advanced Filters.

4.4. Rogue Detection

We propose to perform a set of tasks in order to detect rogue devices in a network, and to do thereafter some actions: First, we propose in ENSec solution, to build which we call “Layer2 Security Database” that contains a full mapping between: Hostname, Manufacturer, Operating System, IP, MAC and some additional information (such as: Layer2 & Layer3 TTL values; the number of detected vulnerabilities) basing on the previous proposed mechanisms and on others implemented on the open source Network Mapper (NMAP). Second, we propose to perform a periodic refresh (for example, each 30 seconds) of all information stored in the previous database in order to detect and to inspect any changes in the network. Third, we propose to reject any information not validated by the Layer2 Security Database (such as ARP replies).

Finally, we propose to implement a mechanism that offers the possibility to detect and to correct default-gateway information using Layer2 Traceroute (to avoid man in the middle device) and Layer3 Traceroute (to assess its ability to route traffic out the network).

Moreover, we propose to classify network devices on three categories, where a set of rules and policies will be applied on each device depending on its category:

- Trusted Devices (TD): are devices that implement the proposed solution and never considered as a suspect during a period of alive-time, (for example, 10 minutes).
- Untrusted Devices (UD): are all other devices even if they implement the proposed solution (before the laps of 10 minutes) or not, but never considered as a suspect.
- Rogue Devices (RD): are all devices even if they implement the proposed solution or not, but considered at least once as a suspect.

We mean by “considered as suspect” all devices doing at least one of the following tasks:

- Changes its MAC address,
- Changes its IP address,
- Replies as a DHCP server and as a default-gateway and found in the middle between devices on the same subnet,

- Replies incorrectly to ARP requests (even about IP and/or MAC),
- Sends invalid Gratuitous ARP,
- Changes the Layer2/Layer3 received TTL value,
- Exceeds the acceptable number & level of detected vulnerabilities,

After a period of time (for example, 600 seconds), we propose to delete all rogue devices from RD-List and include them to UD-List. In parallel, died devices will be detected automatically (using scanning & broadcasting mechanisms) and will be deleted completely from all lists after a period of 180 seconds of dead-state.

4.5. Advanced Filters

The proposed solution offers an advanced way to filter traffic going to and/or from each endpoint device basing on seven parameters:

- Payload content
- Source and Destination Port numbers
- Protocol numbers
- Source and Destination IP addresses
- Source and Destination MAC Addresses
- Ethertype values
- Layer2 and/or Layer3 TTL values

As results, it will be possible to use simple filters using source/destination ports, IP and MAC addresses; but also advanced filters using protocol numbers (such as 1=ICMP, 6=TCP, 17=UDP, 115=L2TP); using Payload content (such as URL, file name or extension); using Ethertype values (such as 0x0806=ARP, 0x8035=RARP); using TTL values which can be helpful when the destination device reside in a well-known location (on the same subnet or on another subnet passing through only one layer3 device for example).

In addition, we propose to use during the filtering process, five methods:

- Exact-Match (EM), for example: Drop traffic from exact (172.16.1.18)
- Less or Equal (LE), for example: Allow traffic where Dst-Port LE (1024)
- Greater or Equal (GE), for example: Allow traffic where Src-Port GE (1024)
- RanGe [from, to] (RG), for example: Allow traffic where TTL RG [10, 255]
- Ignore Part (IgP), for example: Drop traffic from IgP C0-CB-38-**-**-** or *.exe

Therefore, we propose three profiles of filters depending on categories of network devices as described below: allows all traffic to/from trusted devices, allows some traffic to/from untrusted devices, blocks all traffic to/from rogue devices.

5. Discussions

In this section, we prefer discussing some topics and offering answers to related questions in order to clarify all fuzzy points of the proposed framework, and to show its usefulness.

First, the proposed framework is backward-compatible solution with all existing TCP/IP suite protocols, technologies and independent of all hardware and software constraints, which can easily deployed on any device and with any operating system (such as Windows, Linux, Android) regarding its nature as a standalone (portable) application that does not add anything (such new field), but rather modify the interpretation of only one already-existing byte. In addition, endpoints without this solution can easily be discovered and tracked but can never been trusted.

Second, all parameters in the proposed framework, such as timers, categories and advanced filters are customizable; which mean that end-users can tune them according to their needs.

Finally, the proposed framework is not greedy in term of resources, neither in term of CPU/Memory (to store/process the security database's content) nor in term of Bandwidth knowing that sending one ICMP, ARP and DHCP packet per second is less or equal to 500 bytes per second, which represent a negligible value regarding both WAN & LAN links.

6. Conclusion and Perspectives

In this paper we have addressed a new user-centered theoretical framework to improve the security of network devices against most of dangerous attacks (such as man in the middle attacks), especially when connecting to unknown or untrusted networks.

The proposed framework improves considerably the endpoint security against network threats independently of network devices, their software, and also their configuration. It is a standalone, a tunable, a backward-compatible solution with all existing protocols, and finally requires negligible values of resources (CPU, memory and bandwidth).

We plan on implementing it (in the first step) as a java-based standalone (portable) application, where we can deploy it in different endpoints within multiple networks (such as Universities, Hotels, and Internet-Cafes) in order to evaluate its robustness and to perform statistical analysis. We plan also on offering Authentication using Hashing functions and to introduce signature and behavioral Intrusion Prevention System (IPS) that strengthen the endpoint security.

In our opinion, future works in this area should focus particularly on endpoint protection where each device must be responsible about its security, as the number of available and untrusted networks is continuing to grow.

References

[1] D. C. Plummer: An ethernet Address Resolution Protocol. In RFC 826, (1982).

[2] R. Wagner. Address Resolution Protocol Spoofing and Man-in-the-Middle Attacks. The SANS Institute (2001).

[3] Cisco Systems: “Security” In Catalyst 6500 Series Switch Cisco IOS Sofware Configuration Guide, chapter 9, pp. 915-1188, (2007). Available on: http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/ sx_swcg.pdf

[4] Juniper Networks: “Understanding 802.1X, Port Security, and VoIP” in Complete Software Guide for JUNOS® Software for EX-series Switches, chapter 44, pp. 721-760, (2009). Available on: http://www.juniper.net/techpubs/en_US/junos9.3/information-products/topic-collections/ex-series/software-all/book-software-ex-series-93-all.pdf

[5] Y. I. Jerschow, C. Lochert, B. Scheuermann, and M. Mauve, “CLL: A Cryptographic Link Layer for Local Area Networks,” in Proceedings of the 6th Conference on Security and Cryptography for Networks, pp. 21–38, (2008).

[6] Lawrence Berkeley National Laboratory, “ARPwatch Tool” available on: ftp://ftp.ee.lbl.gov/arpwatch.tar.gz (last visit Feb. 16, 2013)

[7] M. Barnaba, Antifork Research Laboratory, “Anticap Kernel Based Patch” available on: https://antifork.org/trac/browser/trunk/anticap (last visit Feb. 16, 2013)

[8] I. Teterin, SecurityFocus Security Community, “Antidote Kernel Based Patch” available on: http://online.securityfocus.com/archive/1/299929 (last visit Feb. 16, 2013)

[9] Danilo Bruschi, Alberto Ornaghi, Emilia Rosti , “S-ARP: a Secure Adderess Resolution Protocol” 19th Annual Computer Security Applications Conference, (2003).

[10] Lootah, W., Enck, W., McDaniel, P.: TARP: Ticket-based Address Resolution Protocol. Computer Networks, vol. 51, no. 15, pp. 4322–4337, (2007).

[11] Seung Yeob Nam; Dongwon Kim; Jeongeun Kim: "Enhanced ARP: Preventing ARP Poisoning-based Man-in-the-Middle Attacks" Communications Letters, IEEE , vol.14, no.2, pp.187-189, (2010).