ENHANCED TECHNIQUES FOR DETECTION AND CLASSIFICATION OF NEIGHBOR DISCOVERY PROTOCOL ANOMALIES

FIRAS (M.H.) S. NAJJAR

UNIVERSITI SAINS MALAYSIA
2016

ENHANCED TECHNIQUES FOR DETECTION AND CLASSIFICATION OF NEIGHBOR DISCOVERY PROTOCOL ANOMALIES

by

FIRAS (M.H.) S. NAJJAR

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

August 2016

DEDICATION

وَقُلْ رَبِّ زِدْنِي عِلْماً

سورة طه اَية (114)

إلى من علمني النجاح والصبر

إلى من افتقده في مواجهة الصعاب

ولم تمهله الدنيا لأرتوي من حنانه..... أبي

وإلى من تتسابق الكلمات لتخرج معبرة عن مكنون ذاتها

من علمتني وعانت الصعاب لأصل إلى ما أنا فيه

التي لم تألُ جهداً في تربيتي وتوجيهي ورعايتي

وعندما تكسوني الهموم أسبح في بحر

حنانها ليخفف من آلامي

.. أمي

اهدي لكم هذا البحث

ACKNOWLEDGEMENT

* * * * * * * * * * **All praises and gratitude to Allah SWT the Almighty** * * * * * * * * *

This modest research will never become true without the contributions from special relatives, friends and colleagues in their own different ways. For this reason, I would like to express my special appreciation and thanks to some of whom it is possible to give particular mention here. First and foremost, all praises and thanks to the Almighty Allah SWT for granting me with patience, guidance and health, as well as giving me the chance to work in an environment such as Universiti Sains Malaysia (USM) and National Advanced IPv6 Center of Excellence (NAv6) particularly. Secondly, I would like to express my sincere and utmost gratitude to my supervisor, Dr. Mohammed M. Kadhum, the one who taught me the meaning of help and support. Special thanks to him for guiding me all the way through in completing this research and giving me the chance to contribute to this field. My gratitude also goes to my field supervisor, Dr. Homam El-Taj, for his support and continuous encouragement in following his path in the academic field. Special thanks to NAv6 director, Prof Dr Rosni Abdullah, and NAv6 management for providing a conductive environment and support during my research. Last but not the least, I would like to thank those who are close to my heart; my beloved mother Raba Enayah for her care, love, du'a and praises, to my wife Rand Enayah , for her endless support and continuous encouragement to finish this research, to my dearest brothers Eng. Ayman Najjar, Ashraf Najjar, and Dr. Mohannad Najjar for making my dream become true, to my precious sisters Dr. Manal Najjar and Dr. Abeer Najjar for teaching me how to persevere and keep motivating me throughout this journey, to all my close friends in Malaysia, Ahmad Slaibih, Mohammed Shehab, and Hamzah, and to all my colleagues and fellow brothers in NAv6: Kamal, Sabri, Nibras, Izanan, Beng, Najjar, Dr. Redwan, Shafeeq, Atheer, Al-Halabi and Dr. Anbar. Finally, I would like to express my thanks to the Unknown Soldier who supports me without notice. I dedicate this work to all of them. **Firas Najjar, Penang, Malaysia, 2016**.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARP	Address Resolution Protocol
CGA	Cryptographically Generated Address
CIP	Current IP address count
CMA	Current MAC Address count
CND	Compact Neighbor Discovery
CPU	Central Processing Unit
CSV	Comma Separated Values
DAD	Duplicate Address Detection
DDoS	Distributed Denial of Service
DHCP	Dynamic Host Configuration Protocol
DMZ	Demilitarized Zone
DoS	Denial of Service Attacks
EFSM	Extended Finite State Machine
FN	False Negative
FP	False Positive
FSM	Finite State Machine
GNS3	Graphical Network Simulator 3
HIDS	Host-based Intrusion Detection

IANA Internet Assign Number Authority

ICMP Internet Control Message Protocol

ICMPv6 Internet Control Message Protocol for IPv6

ID3 Iterative Dichotomiser 3

IDPS Intrusion Detection and Prevention System

IDS Intrusion Detection System

IPS Intrusion Prevention System

IETF Internet Engineering Task Force

IKE Internet Key Exchange

IKE2 Internet Key Exchange 2

INDPMon Intelligent Neighbor Discovery Protocol Monitoring

IOS Internetwork Operating System

IPS Intrusion Prevention system

IPsec Internet Protocol Security

IPv4 Internet Protocol Version 4

IPv6 Internet Protocol Version 6

KNN K-Nearest Neighbor

LAN Local Area Network

MAC Media Access Control

MLD Multicast Listener Discovery

MTU Maximum Transmission Unit

NA Neighbor Advertisement

NA-Count Neighbor Advertisement Count

NAT Network Address Translation

NBA Network Behavior Analysis

ND-Shield Neighbor Discovery Shield

NDP Neighbor Discovery Protocol

NDPMon Neighbor Discovery Protocol Monitor

NIDS Network intrusion detection systems

NIST National Institute of Standards and Technology

NS Neighbor Solicitation

NS-Count Neighbor Solicitation Count

NUD Neighbor Unreachability Detection

OneR One Rule

PART PARtial decision Tree

PKI Public Key Infrastructure

PX network PrefiX

QoS Quality of Service

RA Router Advertisement

RA-Count Router Advertisement Count

RA-Guard	Router Advertisement Guard
RFC	Request For Comments
RIP	Router IP Address
RIPPER	Repeated Incremental Pruning to Produce Error Reduction
RMA	Router MAC Address
ROC	Receiver Operational Curve
RS	Router Solicitation
RS-Count	Router Solicitation Count
RSA	Rivest-Shamir-Adleman cryptosystem
SA	Security Associations
SAVI	Source Address Validation Improvement
SeND	Secure Neighbor Discovery
SLAAC	Stateless Autoconfiguration
SPADE	Statistical Packet Anomaly Detection Engine
SV	Security constants Violation
SVM	Support Vector Machine
TCP	Transmission Control Protocol
TN	True Negative
TP	True Positive
UML	Unified Modelling Language

v6IIDS	Intelligent ICMPv6 DDoS Flooding-attack Detection Framework
VEP	Vulnerability Exploitation Programs
VM	Virtual Machine
ZeroR	Zero Rule

TEKNIK TERTINGKAT UNTUK PENGESANAN DAN

KLASIFIKASI ANOMALI PROTOKOL PENEMUAN JIRAN

ABSTRAK

Kajian ini membentangkan penyelesaian, yang dikenali sebagai "Pemantauan Protokol Penemuan Tetangga Pintar (INDPMon)", berfungsi untuk meningkatkan tahap keselamatan rangkaian IPv6, dengan mengekalkan pengawasan yang berterusan berkenaan insiden Protokol Penemuan Tetangga (NDP), kelemahan, dan kemungkinan serangan dalam membantu keputusan pengurusan risiko organisasi. INDPMon menggunakan pendekatan analisis rangkaian untuk memantau paket lapisan rangkaian, dan menggunakan kaedah protokol stateful untuk menggambarkan anomali protokol dengan tepat. Mesin keadaan terhingga terluas digunakan untuk memahami dan menganalisis tingkah laku dinamik protokol supaya sebarang peristiwa pelanggaran yang menyebabkan anomali NDP dapat dispesifikasi. Peristiwa yang paling diskriminatif dipilih untuk menentukan ciri-ciri set NDP yang akan digunakan untuk menggambarkan kelakuan NDP. Tapak ujian telah digunakan untuk menjana set data NDP dan proses awal prosedur dilakukan kepada set data NDP yang dijana bagi tujuan optimasi. Set data NDP bersama-sama ciri-ciri set NDP digunakan untuk membuat set data perwakilan ciri-ciri NDP yang merupakan tulang belakang INDPMon untuk ramalan dan klasifikasi keputusan. Buat masa ini, alat pemantauan NDP, yang dikenali sebagai NDPMon, adalah penyelesaian yang biasa dinamakan untuk memantau NDP. Walau bagaimanapun, NDPmon menggunakan teknik pemadanan pasif dan bergantung kepada fasa latihan untuk mengenal pasti nod rangkaian yang sah, yang memberi kesan kepada kedinamikan dan skalabiliti. Hasil penilaian menunjukkan bahawa INDPMon mempunyai ketepatan pengesanan yang lebih signifikan berbanding

NDPMon, menjadikan ia satu penyelesaian yang menjanjikan untuk pemantauan NDP. Kecekapan pengesanan membolehkan pentakrifan satu set ciri-ciri NDP yang menggambarkan tingkah laku protokol dengan tepat. Ia adalah penting untuk menyebut bahawa INDPMon hanya boleh mengesan anomali NDP. Oleh itu, ia mesti digabungkan dengan penyelesaian lain untuk membina penyelesaian keselamatan lengkap. Sebagai usaha ke arah memastikan masa depan Internet, sumbangan utama kajian ini ialah memperkenalkan satu rangka kerja yang mampu memantau secara berterusan dan menganalisis proses kelakuan NDP bagi menyediakan sokongan keputusan mengenai pelanggaran piawaian NDP atau polisi organisasi.

ENHANCED TECHNIQUES FOR DETECTION AND CLASSIFICATION OF NEIGHBOR DISCOVERY PROTOCOL ANOMALIES

ABSTRACT

This research presents enhanced solution, called " Intelligent Neighbor Discovery Protocol Monitoring (INDPMon)", for improving the security of IPv6 networks by maintaining constant awareness of Neighbor Discovery Protocol (NDP) incidents, vulnerabilities, and attacks to support organizational risk management decisions. INDPMon adapts a network analysis approach to monitor network layer packets, and utilizes a stateful protocol methodology to precisely describe the protocol anomalies. Extended Finite State Machine is used to understand and analyze the dynamic behavior of the protocol in order to specify the violation events that cause NDP anomalies. The most discriminative events are selected to define the NDP features set which used to characterize the NDP behavior. Testbed has been used to generate NDP dataset and preprocessing procedures are applied to the generated NDP dataset for optimization. NDP dataset along with NDP features set are used to create a representative NDP features dataset which is the backbone of INDPMon for prediction and classifications decisions. Currently, NDP monitoring tool, called (NDPMon), is the commonly cited solution for monitoring NDP. However, NDPMon uses passive matching techniques and depends on training phase to identify network legitimate nodes, which affects the dynamism and scalability. The evaluation results showed that the proposed INDPMon has a significant detection accuracy over NDPMon, which makes it a promising solution for NDP monitoring. The detection efficiency is resulted of defining a set of NDP features that precisely characterizes the protocol behaviors. It is important to mention that INDPMon can only detect NDP anomalies. Hence, it must be

combined with other solutions to build a complete security solution. Working towards securing the future Internet, the major contribution of this research is introducing a framework that is capable of continuously monitoring and analyzing the processes of NDP behaviors to provide decision support regarding the violation of NDP standards or organization policies.

CHAPTER 1

INTRODUCTION

This Chapter provides a clear view of the research presented in this thesis. It discusses the security issues of the Internet and justifies the need for improvement. It includes the research motivation and problem, research objectives, and research scope that are presented in Sections 1.1, 1.2, 1.3, and 1.4, respectively. Section 1.5 presents the research contribution on securing NDP which is done through several steps that are covered in Section 1.6. Finally, the thesis organization is presented in Section 1.7.

1.1 Internet Issues

The massive growth of Internet users led to the exhaustion of Internet Protocol Version 4 (IPv4) (Postel et al., 1981b) address pool that limits IPv4 public addresses to relatively scarce numbers (APNIC, 2011; LACNIC, 2014; RIPE, 2012). Even with the use of Network Address Translation (NAT) (Egevang and Francis, 2001); the rapid growth of Internet usage ensures the exhaustion of IPv4 public addresses at the end (Groat et al., 2011).

To overcome the exhaustion of IPv4 addresses, Internet Assign Number Authority (IANA) had started to allocate addresses using the next generation of addressing Internet Protocol Version 6 (IPv6) (Deering, 1998). IPv6 is expected to replace IPv4, however, IPv4 still carries the majority of Internet traffic; in January 2016, the percentage of users reaching Google services over IPv4 is nearly 90% (Google, 2016). Figure 1.1 illustrates IPv4 pool addresses and the corresponding time spans.

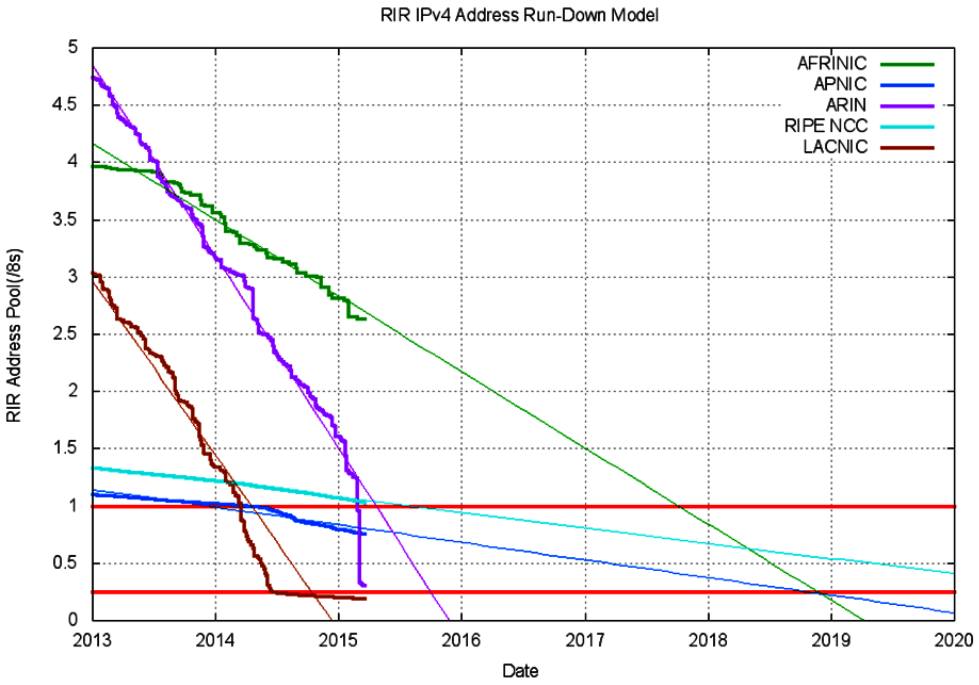

Figure 1.1: Projection of Consumption of Remaining IPv4 Address Pools (Adapted from (Huston, 2015))

Due to the fast exhaustion of IPv4 addresses, organizations are forced to deploy IPv6 and create security policies to deal with the existing security considerations in IPv6 (Najjar and El-Taj, 2015; Saad et al., 2013; Gont, 2013). Although IPv6 design includes the original specification of Internet Protocol Security (IPsec) to secure IPv6, the manual configuration limits its applicability and it is not recommended (Jankiewicz et al., 2011).

As a successor of IPv4, IPv6 was built with security in mind; however, Neighbor Discovery Protocol (NDP) (Narten, Simpson, Nordmark and Soliman, 2007), which is one of the main protocols used by IPv6 suits, has several security concerns (CISCO, 2011). Moreover, the designers of IPv6 suppose that local area network (LAN) consists of trusted users. Hence, NDP trusts every device insides LAN which makes it exposed to various attacks. Also, Eernst and Young (Paul Kessel, 2014), which is a global leader company in assurance, tax, transaction and advisory services, emphasized in their survey in 2013 and 2014 that employees were seen as the most likely source of an attack, and still seen as a significant risk. Furthermore, most

of LANs involve a security privileges hierarchy between users and these privileges must be granted between users.

In addition, IPv6 designers have not anticipated the large deployment of wireless networks which has a minimal or no link-layer authentication, such as public networks in airports and other places. Likewise, they suppose that LANs can be physically protected (Arkko et al., 2002).

Therefore, in communication systems, the use of new technologies which have security limitation due the lack of registration and authentication makes these systems prone to attacks. In IPv6 networks, NDP allow the connected devices to generate their IP addresses and start to communicate with other devices without any registration or authentication inside the network. In addition, most of the existing operating systems are IPv6 enabled by default, thus, LANs are exposed to IPv6 attacks even if IPv6 is not the protocol that is used for communication. Therefore, the only way to prevent these attacks by manually disabling IPv6.

The differences between IPv4 and IPv6 change the types of attacks. The main change in IPv6 is on how IP interacts with the link layer; for example, NDP replaces Address Resolution Protocol (ARP) (Plummer, 1982), thus; Internet Control Message Protocol for IPv6 (ICMPv6) (Conta and Gupta, 2006) messages is used in address resolution instead of Internet Control Message Protocol (ICMP) messages(Postel et al., 1981a). Moreover, in IPv4 network, address resolution attacks can be prevented by disabling some ICMP messages, however, in IPv6 network, disabling ICMPv6 messages affect the network, therefore, techniques such as SEcure Neighbor Discovery (SEND) is a must to secure NDP (Arkko et al., 2005). Supriyanto et al. (2013) stated *"Without SEND, we will fall prey to the same class of attacks we faced in IPv4 over networks"*. The next subsection highlight the effects of using IPv6 on common attacks to confirm the security implications of using IPv6.

IPv6 Commonly Known Attacks

Reconnaissance Attack: Reconnaissance attack seems more complicated in IPv6 networks since the usual subnet size is 64 bits and with the same speed of scanning IPv4 subnet, it would take 60 billion years to scan all addresses, which makes regular scanning techniques impossible unless an adversary uses different approaches; however, Chown (2008) mentioned that some techniques reduce the subset size, as if the adversary knows the Ethernet vendor prefix, which results in reducing the search space to 48 bits and if the Ethernet vendor prefix is known , the search space may reduce to 24 bits. Moreover, Network Mapper tool has the ability to collect all IPv6 messages by sending only one echo message (Lyon, 2009).

ARP Attacks. In IPv4, ARP spoofing enables adversary parties to intercept, modify, or even stop data in transit. In order to prevent ARP spoofing, Cisco implemented a new technique called (snooping) (Cisco, 2012). On the other hand , the situation is considerably different in IPv6; even though ARP is replaced by NDP, similar attacks are still possible through Neighbor Solicitation and Neighbor Advertisement spoofing (Nikander et al., 2004).

Smurf Attack. In IPv4-Smurf attacks, the adversary sends an echo-request message (ping utility) with a destination address of a subnet broadcast and a spoofed source address using the host IP address of the victim. This derives all the devices on the subnet to respond to the spoofed source IP address, and consequently, flooding the victim with echo-reply messages. In IPv6, the concept of an IP broadcast is removed; however, the Smurf attack can be performed using the generated Neighbor Solicitation and Neighbor Advertisement messages (Parameter Problem ICMPv6 message) (Scott Hogg, 2009).

Flooding attack. Flooding attack is one of the most frequent attacks present in IPv4 networks. Also, this type of attack affects IPv6 networks by flooding the network of NDP mes-

sages (Chown and Venaas, 2011).

Application Layer Attack. The only change in Application-layer attack by applying IPv6 is the propagation of worms (Chen et al., 2003; Kamra et al., 2005). Traditionally, worms make local and wild scanning to find victim hosts (Zou et al., 2005), which makes it unlikely to succeed in IPv6 environment; but as mentioned earlier, while taking advantage of local knowledge and patterns in address space assignment, the attack program can reduce the search space considerably. Also the use of all nodes multicast address.

1.2 Research Motivation

NDP uses stateless mechanisms without any authentication or registration, therefore, NDP nodes are exposed to attacks, and because most operating systems of the network computers are IPv6 enabled by default, this makes IPv4 nodes also exposed to the earlier presented NDP attacks until IPv6 is manually disabled. Most of these problems are inherited from the old version IPv4. However, attacks in IPv4 can be prevented by disabling some protocol messages while these messages are the core of the IPv6 network. Disabling them affects IPv6 main functions, which makes it an impossible choice.

In IPv6 networks, NDP allows any connected node to generate its own IP addresses and start communication with other connected devices without registration or authentication inside the network. Consequently, the attacker can claim himself as any device inside LAN by spoofing the protocol messages. For instance, Figure 1.2 illustrates the router messages spoofing attack.

In this scenario, the attacker sends spoofed messages to all network hosts inside the network, advertising himself as the default router. Any host receives these messages, immediately updates its default router to be the attacker node without any verification, and starts to commu-

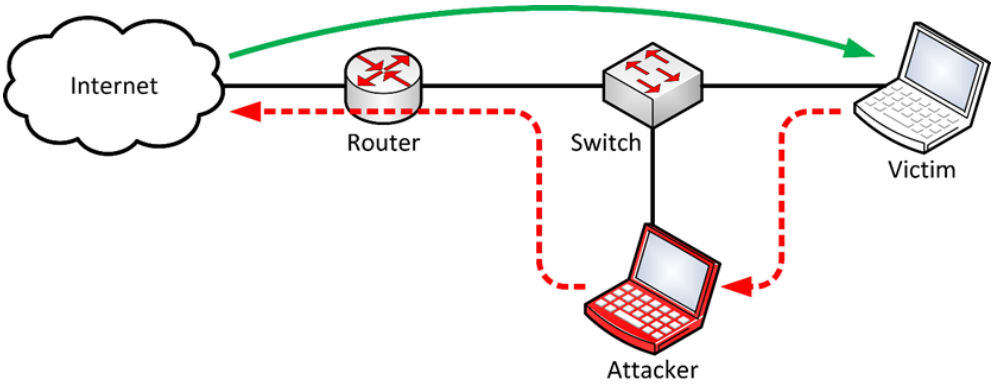

Figure 1.2: Router spoofing Messages(Adapted from (Weber, 2013))

nicate with the fake router as the default router. Therefore, the attacker can launch MITMA by capturing,sniffing, and forwarding the victim's traffic, or launch Denial of Service (DoS) attack by sending spoofing messages containing the address of unreachable router.

Figure 1.3 illustrates another scenario where the attacker can spoof a cretin Neighbor Advertisement message to launch MITMA attacks or DoS attacks. Through spoofing the host messages, the attacker redirects all network traffic over his node, and starts capturing, sniffing, and even changing redirected traffic.

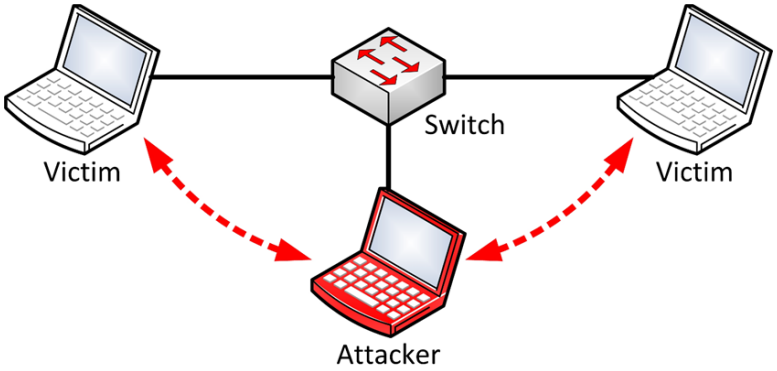

Figure 1.3: Neighbor Advertisement spoofing Messages(Adapted from (Weber, 2013))

Moreover, the attacker can launch DoS attacks by answering all messages generating by duplicated address detection algorithm which is used to verify the uniqueness of the generated IP addresses. Answering these messages prevents the new hosts from getting new addresses; thus, the host will not be able to access the network. Furthermore, the attacker can flood the

entire LAN with thousands of fake router addresses, keeping the hosts busy in generating new global IP addresses, which exhausts the host's resources and eventually freezing them totally.

In order to secure NDP, many solutions were proposed to prevent or detect NDP anomalies. Most of the organizations employ the latest technologies in Intrusion detection and prevention systems (IDPS), which monitor all events occurring in computer systems or networks, analyzing them for signs of possible incidents and attempting to stop detected possible incidents. However, detecting every intrusion attempts increases the number of false positive alerts (notifying the administrator of normal action as anomaly action). In a large organization, usually IDPs generate thousand to millions of alerts per day; most of them are false positive, which means more malicious events are detected. However, more analysis resources are needed to differentiate false positives from true malicious events. This large number of alerts put system administrators in critical situations, where all alerts must be scanned and classified as normal or threats, then they must response promptly to stop threats or slow down them, which is a very tough mission.

Attack prevention solutions such as IPsec intends to secure IPv6 network; yet, it is prone to bootstrapping and needs manual configurations which make it limited to small networks. Moreover, it must verify the ownership of dynamic IP addresses which impossible (Jara et al., 2014; Rantos et al., 2013). Another prevention solution is SEND which it's the best solution where IPsec found to be an impractical choice, however, using public key infrastructure for generating addresses increases the protocol complexity and need a high computational processed which make it prone to DoS. Moreover, statically configured addresses and addresses generated using fixed identifiers can not be protected by SEND (Frankel et al., 2010). Trust-ND, ND-Shield, SAVI, and RA-Guard are another examples of preventing solutions, most of them violate the design principle of NDP in terms of complexity and overhead, or needs a specific hardware with manual configuration. In this regard, Narten, Thomson and Jinmei (2007)) stated that:

"stateless autoconfiguration mechanism requires no manual configuration of hosts, minimal (if any) configuration of routers, and no additional servers".On the other hand, monitoring solutions do not need any modification to the protocol, and they provide a constant awareness of protocol incidents, which make them essential for securing NDP network.

1.3 Research Problem

Existing passive monitoring solutions such as NDPMon use passive training phase to define legitimate nodes, which lack the scalability and dynamism of the solutions. In the training phase, all connected devices are considered legitimate even they are compromised, which make it difficult task to detect the compromised nodes. Moreover, most operating systems by default use multiple IP addresses for privacy reasons (Chown and Venaas, 2011). Using passive matching monitoring solution in such environments is impossible since it must run the passive training every time new address is added to the new address to legitimate address list. This makes it one of the main challenges facing these solutions and makes it limited to specific environments. For instance, public environment, each time a new station connect or leave makes passive matching techniques are useless by increasing the numbers of false positive alarms. Also existing passive monitoring solutions do not have any prediction or classification capabilities, which make them failed in detecting novel attacks.

To avoid passive training phase, NDP active monitoring solutions generate probes into the network for further analysis. These probes increase the network overhead, and could be used in Smurf attacks. Moreover, it can not detect the attacks generated from legitimate nodes. Prio-drop, Host Based IDS, and v6IIDs are another examples of monitoring solution. Most of them use passive matching techniques which generate a high rate of false positive alerts, which affects the solution trust, and consequently, systems administrators detect most of the attacks after they have already happened (Chauhan et al., 2012; Garcia-Teodoro et al., 2009).

Furthermore, one of the major problems that disrupts the research on securing NDP using intelligent techniques is the absence of representative NDP dataset. Most of existing benchmark dataset do not have IPv6 packet (Najjar and Kadhum, 2015), and others do not have all features that accurately characterize the protocol behaviors (SAAD et al., 2014).

1.4 Research Goal and Objectives

The ultimate goal of this research is to build a solution that improves the IPv6 network security without increase in the complexity of NDP or adding overhead to the network by monitoring NDP main processes, and report any violation to the protocol standards. To achieve the research goal, the following objectives have to be accomplished:

1. To design a solution that improved of accuracy in detection and classification of NDP anomalies, which would help in better decision making to secure the network environment. Therefore, this objective has been further broken down into the following sub-objectives:

    - To model the NDP behaviors to provide a better visualization, understanding, and assist in specifying the behaviors of the protocol.
    - To provide the prediction and classification capabilities to the proposed solution by defining an informative NDP features set that accurately characterizes the normal and abnormal protocol behaviors. These generated features are the backbone of the solution prediction and classification capabilities.
    - To create representative NDP dataset that assist in generating NDP features dataset that is used to train the proposed solution.

2. To verify and validate the proposed solution design in terms of detection accuracy, sensitivity, and specificity, and compare it to the existing NDP detection approach in order

to determine the proposed solution effectiveness at securing NDP.

1.5 Research Scope

This research concentrates on improving the security of IPv6 network by maintaining monitoring and analysis capability of NDP processes to provide decision support regarding situational awareness and a violation of expectations. Therefore, research focuses on detecting NDP anomalies which occur only inside IPv6 LAN. NDP uses five ICMPv6 messages to complete its functions. Accordingly, the proposed solution only inspects packets that contain these five messages, since these messages are the main source for defining the features used to detect NDP anomalies. Due the lacking of IPv6 dataset and resources, Testbed includes seven nodes were used to create NDP dataset, and also used to verify and validate the proposed solution and the proposed solution evaluated only according to the correctness of NDP anomalies detection and classification. Figure 1.4 illustrate the network scope of the research. The area inside the red shape is the network scope interest of this research.

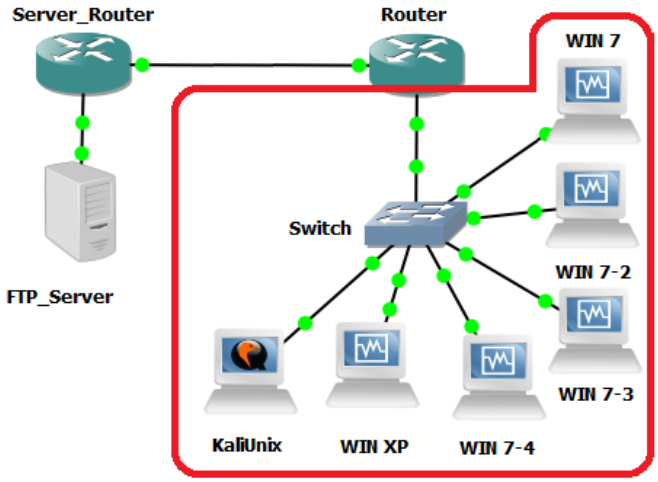

Figure 1.4: Research Scope

1.6 Research Contributions

IPv6 is considered as the backbone of the future Internet; however, it is prone to attack since NDP uses unsecure mechanisms. Thus, the main contribution of the research presented in this thesis is the enhancement of the overall IPv6 security environments where the trust between nodes is a critical issue. In regard to the importance of contribution towards securing the future Internet, an intelligent solution is proposed, which accurately detects any signs of possible NDP incidents that are considered violations or imminent threats of violation to organizations security policies. In terms of the contributions to the body of knowledge, extended finite state machine modelling is proposed for NDP behavior study and analysis, where it can be of great benefits for further research on dynamics of NDP in order to build security tools for NDP. Such modelling can provide informative NDP features set that can be defined to accurately characterize the protocol behaviors. These features describe the behaviors of NDP and help in addressing the protocol violations. In addition, this research provides a baseline of how the machine learning concept can be adapted to predict novel attacks on NDP and helps in making intelligent decisions to stop or deal with the protocol threats. Also, utilizing the machine learning concept helps in taking the advantage of the optimal NDP features set for defining potentially predictive relationships used to detect new attacks.

1.7 Research Steps

Figure 1.5 illustrates the research steps followed in conducting this research:

Step 1: Literature Review. This phase covers the background of NDP and its main functions in addition to a detailed analysis of the NDP weakness and threats. Also, approaches that have been developed to secure NDP and most techniques used to detect NDP attack are presented.

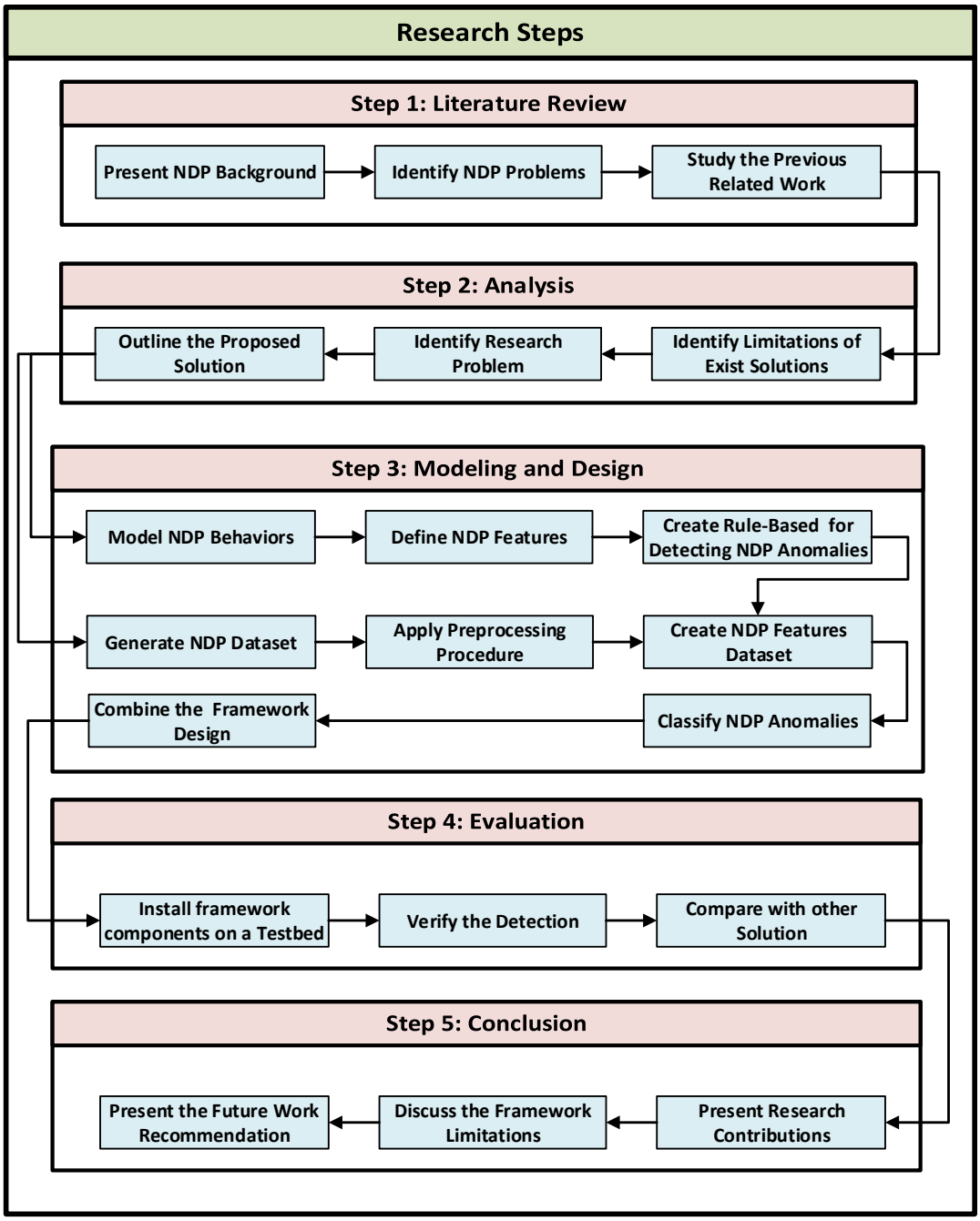

Figure 1.5: Research Steps

Step 2: Literature Analysis. This phase discusses and analyzes the major approaches used to secure NDP, and identifies advantages and limitations for each approach. Hence, it provides a better understanding of current solutions limitations, research problem and scope, which gives the knowledge to outline the proposed solution.

Step 3: Modelling and Design. This phase discusses the modelling of NDP normal be-

havior and utilizes stateful anomaly detection method to define the appropriate NDP feature set which in turn is used to accurately detect NDP anomalies. In addition, in this phase, NDP dataset generation is discussed along with a description of the tools that are used in the generation and preprocessing procedures. The generated dataset along with the defined NDP features are used to create the NDP features dataset in order to utilize a machine learning technique to classify the behavior of NDP. Anomalies classification helps in understanding the NDP network behaviors. The combined proposed solution component design is the output of this stage.

Step 4: Evaluation. In this phase, a real world case study is used to evaluate the efficiency of the proposed solution in terms of detection correctness. Different NDP behavioral scenarios are performed and discussed to generate normal and anomalies behavior. Finally, the proposed solution is validated by comparing it with another existing solution in terms of detection accuracy and to ensure its usefulness.

Phase 5: Conclusion. This phase summarizes the research work, presents contributions, highlights the limitations, and suggests some future work.

1.8 Research Organization

The rest of the thesis is organized as follows:

Chapter Two covers the detailed background of NDP along with fundamental concepts related to this research, as well as, the current existing related work and highlight the limitation of each solution.

Chapter Three details out the proposed solution methodology. It explains the proposed solution requirements and presents the sui table methods and approaches to meet these requirements.

Chapter Four introduces the new solution that is called "Intelligent Neighbor Discovery Protocol Monitor (INDPMon)", for detecting and classifying NDP anomalies. This chapter describes the structural design and implementation of INDPMon.

Chapter Five presents a detailed performance evaluation of the proposed INDPMon solution, in comparison with the performance of the common existing solutions, namely NDPmon.

Chapter Six summarizes the findings of the research presented in this thesis and provides recommendations for further research work.

CHAPTER 2

LITERATURE REVIEW

This chapter presents a detailed and ncomprehensive background of NDP and the fundamental concepts related to this research. It discusses the related works on securing NDP and highlights the limitations of each solution, enwhich provide the motivation to this research. This chapter is organized as follows: Section 2.1 provides an overview on NDP, Section 2.2 discusses Intrusion Detection and Prevention systems, Section 2.3 demonstrates the related works, and finally the chapter is summarized in Section 2.4.

2.1 NDP Overview

NDP is a set of messages and processes which determines relationships between nodes (routers and hosts) and their neighbor's in the same network. NDP replaces some protocols used in IPv4 such as Router Discovery, Address Resolution Protocol (ARP), Internet Control Message Protocol (ICMP), and ICMP Redirect. Additionally, NDP provides new functionalities that are used for Duplicate Address Detection (DAD) and Neighbor Unreachability Detection (NUD). IPv6 NDP allows nodes to identify their neighbors on the same LAN and advertise their existence to other neighbors.

One of the most interesting and potentially valuable addressing feature implemented in IPv6 is the ability of all nodes to automatically configure their addresses using Stateless Address Autoconfiguration (SLAAC) (Narten, Thomson, et al., 2007) without using any stateful configuration protocol, such as Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Bound et al. (2003). The use of SLAAC gives the ability to IPv6 host to generate link-local

and global addresses without any manual intervention. The behavior of IPv6 NDP consists of the following processes:

- Generating IP addresses using SLAAC.
- Verifying the uniqueness of the generated IP address using DAD.
- Resolving IP addresses to their link-layer addresses.
- Maintaining reachability track with neighbor nodes using NUD.
- Using Redirect Message for advertising a better next-hop

2.1.1 NDP Messages

To complete NDP processes, NDP uses five ICMPv6 messages, which are:

1. Router Advertisement (RA) messages are originated by routers and sent periodicity, or sent in response to Router Solicitation to advertise their presence and send parameters such as Router Prefix, Maximum Transmit Unit (MTU), a list of prefix, and hop limits. Furthermore, RA contains the information on how a host configures its global IP address using stateless of stateful technique.
2. Router Solicitation (RS) messages are originated by hosts at the system startup to request for an immediately RA message rather than waiting for the next scheduler timer.
3. Neighbor Solicitation (NS) messages are originated by hosts that attempt to discover the link-layer addresses of other nodes on the same local link, or originated by DAD algorithm, or originated to verifying the reachability of a neighbor.
4. Neighbor Advertisement (NA) messages are sent to advertise the changes of a host MAC address and IP address, or in response to NS messages which could be sent by address

resolution process, or by NUD process. Three oodflags are included in NA message: Router flag, when it is set, it indicates that the sender is a router; Solicited flag indicates that the NA was sent in response to NS message; and Override flag which indicates that the NA message information should update an existing cache entry.

5. Redirect messages are used to redirect traffic from one router to another.

2.1.1(a) NDP Messages Options

NDP messages include zero or more option depending on the information sent. There are five types of options as shown in Table 2.1.

Table 2.1: NDP Messages Options

Type	Name	Description
1	Source Link-layer	It is used by RA, RS, and NS to present the Link-layer address of the packet sender.
2	Target Link-layer	It is used by RA and NA to present the address of target link-layer address.
3	Prefix Information	It provides hosts with on-link prefixes and prefixes for Address Autoconfiguration.
4	Redirected Header	It is used by the redirect messages to contain all or part of the redirected packets.
5	MTU	It is used by RA to inform all nodes to use same MTU value in the same link.

2.1.1(b) NDP Message Validity Checks

All NDP message must satisfy the following conditions; if any violation is detected, the message must be discarded (Narten, Simpson, Nordmark and Soliman, 2007):

- Hop Limit field equals 255 since all NDP messages must not be forwarded by a router.
- Valid ICMPv6 message checksum.
- ICMPv6 message code is equal to zero.

- Target address must not be a multicast address.
- Message options must have a length greater than zero.

Neighbor Advertisement message validation has additional validation checks which are as follows:

- Solicited flag must be zero if the destination IP address is a multicast address.
- ICMPv6 length equals 24 octets or more.

Neighbor Solicitation message validation has additional validation checks which are as follows:

- If the source IP address in a NS message contains an unspecified address, the destination IP address of the NS message must be a solicited-node multicast address, and there is no source link-layer option in the message.
- ICMPv6 length equals 24 octets or more.

Router Advertisement message validation has additional validation checks which are as follows:

- The source IP address of the RA message must be a link-local address.
- ICMPv6 length equals 16 octets or more.

Router Solicitation message validation has additional validation checks which are as follows:

- If the source IP address of the RS message is an unspecified address, the source link-layer address option of the NS message must not be included in the message.
- ICMPv6 length equals eight octets or more.

2.1.1(c) NDP Constants

NDP constants used to regulate nodes NDP message generation, and allows NDP to operate over links with widely varying performance characteristics. Table 2.2 illustrates the values and description for each constant.

Table 2.2: NDP Constants

NDP Constant Name	Constant Value	Description
Router constants		
MAX_INITIAL_RTR_ADVERT_INTERVAL	16 seconds	Maximum time between unsolicited RA
MAX_INITIAL_RTR_ADVERTISEMENTS	3 transmissions	Maximum number of unsolicited RA
MAX_FINAL_RTR_ADVERTISEMENTS	3 transmissions	Maximum number of RA before shutdown
MIN_DELAY_BETWEEN_RAS	3 seconds	Minimum time between RA
MAX_RA_DELAY_TIME	0.5 second	Maximum delay time for RS response
Host constants		
MAX_RTR_SOLICITATION_DELAY	1 second	Maximum delay time for the host to send RS
RTR_SOLICITATION_INTERVAL	4 seconds	Maximum time between RS
MAX_RTR_SOLICITATION	3 transmissions	Maximum RS for each
Node constants		
MAX_MULTICAST_SOLICIT	3 transmissions	Maximum multicast NS for address resolution
MAX_UNICAST_SOLICIT	3 transmissions	Maximum number of NS solicitation
MAX_ANYCAST_DELAY_TIME	1 second	Maximum sending delay for anycast
MAX_NEIGHBOR_ADVERTISEMENTS	3 transmissions	Maximum number of NA
RetransTimer	1,000 miliseconds	Retransmission time for NS or NA
AdvValidLifeTime	Default 30 days	Prefix valid Life-Time
AdvPreferredLifeTime	Default 7 days	Prefix preferred lifetime
DELAY_FIRST_PROBE_TIME	5 seconds	Time waited until upper protocol confirmation
DupAddrDetectTransmits	Default 1, 0 disable	Maximum number of NS for DAD algorithm

2.1.2 Stateless Autoconfiguration

Stateless Autoconfiguration processes include the generation of link-local IP address, generation of global IP address, and the verification of the uniqueness of the generated addresses using DAD algorithm. While configuring its IP address, each node moves through different IP states which are as follows: Tentative Address state, Preferred Address state, Deprecated

Address state, Invalid Address state, and DAD-Failure state.

All IPv6 nodes must have a link-local address in order to (i) determine the link-local addresses of their neighbors, (ii) to communicate with routers, and (iii) to maintain reachability information for communicated nodes (Carpenter and Jiang, 2014). Failure of configuring link-local addresses violates one of the design goals of SLAAC mechanism. Consequently, the process will totally fail and the nodes become blind; yet they cannot communicate with each other, particularly in the absence of routers and DHCPv6. Manual intervention must be done to reconfigure the nodes.

2.1.2(a) Link-Local Address Generation

On system startup, each IPv6 node automatically creates a link-local address using the EUI-64 method or using other methods such as random generation (Hinden and Deering, 2006). Before assigning link-local address to an interface, the node enters into Tentative Address state until the uniqueness verification of the tentative address is done using DAD (or enhanced DAD (Asati et al., 2015)) algorithm which is an enhancement to DAD algorithm to automate the detection of looped back NDP messages used by DAD, by adding new option variable called Nonce). DAD and enhanced-DAD abbreviations will be used interchangeably in this research.

While nodes are in Tentative Address state, their interfaces discard all received packets except DAD related packets. Tentative Address uniqueness verification performed by DAD algorithm using the information provided in NS and NA messages. The process of DAD involves:

1. On system startup, each node joins the link-local multicast group for all nodes, as well as a solicited-node group for tentative address, in order to send and receive DAD packets. Afterward, the node sends NS carrying the host tentative address with random generated

Nonce, to the solicited-multicast group and waits for a designated NA.

2. If no NA is received within the specified retransmit time, or receiving NS with the same Nonce number, the tentative address is considered unique within its neighbors, and the node address state becomes preferred state. The leased time for the unique link-local address is infinite length time.

3. If NA is received within the specified retransmit time, and the target field contains the tentative address or receiving NS with target field that contains the tentative address with different Nonce number, or receiving NS with source address field that contains the tentative address, then, the tentative address is considered invalid, and the node address state becomes DAD-Failure state. Thereafter, the node leaves the solicited-node multicast group if it does not have another IP address has the same solicited-node multicast group with the tentative address. In case that a duplicated link-local is generated from hardware identifier, the IP operation will be disabled; otherwise, the IP operation on the interface continues Narten, Simpson, Nordmark and Soliman (2007).

The nodes stay in Preferred Address state while the interface is enabled; otherwise, whenever the interface becomes disabled manually, or disabled by hardware failure, or the interface cable unplugged, the nodes restart the procedure of configuring the link-local address. Sometimes the processes of DAD are not reliable; mainly when there are two hosts performing DAD processes at the same time. Under this circumstance, the two hosts may have the same address (Narten, Simpson, Nordmark and Soliman, 2007). DAD process must be performed on all unicast addresses (Asati et al., 2015).

2.1.2(b) Global Address Generation

In order to configure the IPv6 global address, hosts must communicate with a router or DHCPv6 to obtain the prefix information options that will be appended to the interface identifier (Singh et al., 2010). This information is sent via RA message or DHCPv6. Moreover, RA provides information that guides the host on how to perform address configuration using a stateless or stateful technique.

In addition, RA provides important IPv6 parameters such as default route, MTU, default Hop-limit, and Domain Name Server option (Jeong et al., 2010). The process of global Addresses creation is as follows:

1. On system startup, in the absence of RA, the host sends RS to routers multicast group address.
2. If no RA is received within the specified retransmit time, the host keeps sending NS messages every BackOff timer (Krishnan et al., 2015).
3. If RA is received, then the host generates global address by appending the received prefix to the interface identifier; therewith, the host enters Tentative Address state, along with proceeding DAD process. In case of DAD process failure, the state of the address becomes DAD-Failure state. Otherwise, the process of DAD succeeds and each address must have two timers derived from RA, namely *PreferredLifeTime* and *ValidLifeTime*. Whenever the host has a global address, *PreferredLifeTime* and *ValidLifeTime* are refreshed through RA.

Routers must periodically send RA messages announcing their availability and information about the network. Whenever the host receives these packets, it must update its related entries, such as default router. RA must only refresh addresses derived from a prefix option advertised

in RA. Furthermore, the valid time of the prefix must be greater than preferred time of it. If the host's IPv6 global address preferred time is greater than the it's valid time, the host enters the Preferred Address state. Whenever the preferred time becomes zero, and the valid time is greater than zero, the host enters Deprecated Address state; otherwise, the IP address becomes invalid and the host enters the Invalid Address state. While the host is in Deprecated Address state, the address should only be used by applications that have been using it until the session is finished, in order to not disrupt the connection.

2.1.3 Address Resolution and Neighbor Unreachability Detection

NDP replaces the ARP for resolving node's addresses to MAC address. NDP uses Address Resolution process to resolve IPv6 addresses and uses NUD to keep track of the reachability to other nodes. In Address Resolution and NUD processes, NDP uses NS and NA messages.

With the aim of sending data to neighbor nodes, each node must know the MAC address of the corresponding neighboring node. If the node does not have the MAC address of the receiver, it initiates the Address Resolution process by sending NS to the solicited-node multicast group of the receiver address requesting its MAC address. Meanwhile, the sending node creates a record for the receiver IP address in its cache and makes the address status Incomplete.

Whenever the sender node receives a solicited-NA message, which is a response to NS message, the sender node updates the cache record with the MAC address and changes the address status to Reachable for a specific time; after the time expired, the status of the address becomes Stale. Moreover, the status of the address becomes Stale at any time the host receives unsolicited NA, NS, or Redirect messages; and updating the cache record with new MAC address as well. When a node sends packets to Stale address, the address status becomes Delay for *DELAY_FIRST_PROBE_TIME* (in seconds) to give the upper protocol the chance to

update the address reachability. Hence, the address status becomes Reachable when the node receives a confirmation from the upper protocol. If *DELAY_FIRST_PROBE_TIME* is expired, the status of the address becomes Probe. In Probe state, the nodes send a unicast NS message to confirm the reachability of the address. If there is no confirmation received, the address status becomes Unreachable and the node sends Solicited-Multicast NS every Exponential backoff timer (Gashinsky and Nordmark, 2014).

2.1.4 Redirect Messages

NDP Redirect messages are similar to ICMP redirect, which are sent to hosts by routers to inform them of a better next-hop destination. Redirect message assists hosts to choose the most efficient routing path. Once the host receives a redirect message, the host either updates the existing cache record for the target or creates a new cache record.

2.1.5 NDP Common Attacks

Nikander et al. (2004) specify three different trust models for securing NDP, each model has different level of trust:

- High trust; all authenticated nodes in this model trust each other and treat each IP layer message as a legitimate one. In this model, all NDP message do not contain any false information. Corporate Intranet Model is an example of this model.
- Semi trust; where there are some nodes are trusted inside the network. For example, legitimate routers that are considered the only trusted nodes which routes all packets between the local network nodes and any connected external networks. Public Wireless Network with an operator is an example of this model.
- No trust, where there is no trust between nodes at the IP layer. Ad hoc network is an