الجـــمـهـوريـــة الجـزائـريـة الديمـقـراطيـة الـشـعـبيـة

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

وزارة التـــعـليـــم العـــــالي والبـــحث العـــــلـمـــي

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

جـــــــــــامعة سعيدة – د. الطاهر مولاي –

University of Saida - Dr. MOULAY TAHAR
Faculty of Technology

A Dissertation Submitted to the Department of Telecommunications in Partial Fulfilment of the Requirements for Degree of Master of
Networks and Telecommunications

Presented by:

BENAISSA Mohamed El Mehdi Seif Eddine

HAMRI Abd Elkader Iheb

Detecting and Predicting Massive Cyber Attacks Using Adaptive EWMA Based Models

Defended on June, 23th 2024 in front of the jury composed of:

Mr. OUARDI Aissa	MCA	President
Mr. BOUYEDDOU Benamar	MCA	Supervisor
Mme. BOUCHENAK Sofia	MAA	Examiner

2023 / 2024

Dedication

We dedicate this work to our beloved parents, whose endless support and encouragement have been our source of strength and inspiration. Their unwavering love has guided us through every challenge.

We also extend our deepest gratitude to our supportive siblings, friends, mentors, and classmates, whose advice and encouragement have been invaluable in completing this journey. We will forever cherish their kindness and generosity.

Acknowledgements

All the praises and thanks be to Allah Almighty, the Giver of bountiful blessings and gifts.

Prayers and peace of Allah be upon the noble Prophet and upon his family and companions.

Our deepest gratitude goes to Mr. BOUYEDDOU Benamar for his invaluable guidance, which significantly contributed to the success of our research.

We are profoundly thankful to our families for their unwavering support, and to our classmates for their valuable advice, ensuring the timely completion of this project.

Finelly, we appreciate the examination jury members for their time and feedback, which have been integral to our academic development.

Contents

Contents

Contents

List of Figures

List of Figures

List of Figures

List of Tables

List of Abbreviations

ACK Acknowledgment

CL Center Line

CUSUM Cumulative SUM

DARPA Defense Advanced Research Projects Agency

DARPA99 DARPA Intrusion Detection Evaluation dataset of 1999

DDoS Distributed Denial of Service

DoS Denial of Service

EWMA Exponentially Weighted Moving Average

AEWMA Adaptive Exponentially Weighted Moving Average

D-EWMA Double Exponentially Weighted Moving Average

T-EWMA Triple Exponentially Weighted Moving Average

ICMP Internet Control Message Protocol

IP Internet Protocol

LCL Lower Control Limit

SPC Statistical Process Control

TQM Total Quality Management

SYN Synchronization

TCP Transmission Control Protocol

TCP/IP Transmission Control Protocol/Internet Protocol

UCL Upper Control Limit

UDP User Datagram Protocol

FSS Fixed sample size

VSS Variable sample size

VSSILF Variable sample size Integer Linear Function

RIP Routing Information protocol

HTTP Hyper Text Transfer Protocol

FTP File Transfer Protocol

OSPF Open Shortest Path First

SMTP Simple Mail Transfer Protocol

Abstract

In our digitally interconnected world, the prevalence of cyber-attacks poses significant threats to network security and integrity. Among these, Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks stand out for their potential to cause substantial financial harm and reputational damage. Control charts, statistical tools designed to monitor process variations, offer a promising way for detecting and mitigating such attacks. This dissertation investigates control chart's efficiency in identifying DoS and DDoS attacks through network traffic analysis. Here, we examine the performance of EWMA, VSS-EWMA, VSSILF-EWMA, and ASW-EWMA under various attack's scenarios when evaluating their capabilities using the DARPA99 dataset. Our findings indicate that these control charts can effectively detect DoS and DDoS attacks, provide enhanced sensitivity and reduce false alarms rates.

Keywords: TCP/IP, Cyber-Attacks, DoS/DDoS attacks, TCP SYN Attacks, Smurf Attacks, Control charts, EWMA chart, adaptive EWMA chart, DARPA99 dataset.

ملخص

في عالمنا الرقمي المترابط، يشكل انتشار الهجمات الإلكترونية تهديدات كبيرة لأمن الشبكات وسلامتها. من بين هذه الهجمات، تبرز هجمات حجب الخدمة (DoS) وهجمات حجب الخدمة الموزعة (DDoS) لقدرتها على التسبب في أضرار كبيرة وإلحاق الضرر بالسمعة. توفر مخططات التحكم، وهي أدوات إحصائية مصممة لمراقبة تقلبات العمليات، مسارًا واعدًا لاكتشاف هذه الهجمات والتخفيف منها. تبحث هذه الأطروحة في فعالية مخططات التحكم في تحديد هجمات DoSو DDoSمن خلال تحليل تدفق بيانات الشبكة. نفحص أداء مخططات التحكم EWMA و VSS-EWMA و VSSILF-EWMAو ASW-EWMA، من خلال تقييم قدراتها تحت سيناريوهات مختلفة من الهجمات الإلكترونية باستخدام مجموعة بياناتDARPA99. تشير نتائجنا إلى أن هذه المخططات تستطيع أن تكتشف بفعالية هجمات DoS وDDoS، تساهم في تحسين الحساسية والتقليل من الإنذارات الكاذبة .

الكلمات المفتاحية :TCP/IP ,الهجمات الإلكترونية، هجماتDos/DDoS ، هجماتTCP SYN ، مخططات التحكم، ، مخططEWMA ، مخطط EWMA التكيفي، مجموعة بياناتDARPA9 .

Résumé

Dans notre monde numériquement interconnecté, la prévalence des cyberattaques constitue une menace significative pour la sécurité et l'intégrité des réseaux. Parmi celles-ci, les attaques par déni de service (DoS) et par déni de service distribuée (DDoS) se distinguent par leur potentiel à causer des dommages substantiels et à nuire à la réputation. Les cartes de contrôle, outils statistiques conçus pour surveiller les variations des processus, offrent un axe prometteur pour détecter et atténuer ces attaques. Ce travail étudie l'efficacité des cartes de contrôle dans l'identification des attaques DoS et DDoS à travers l'analyse du trafic réseau. Ici, nous examinons les performances des cartes de contrôle EWMA, VSS-EWMA, VSSILF-EWMA, et ASW-EWMA sous différents scénarios de cyber-attaques en évaluant leurs capacités en utilisant la base de données DARPA99. Nos résultats indiquent que ces cartes de contrôle peuvent détecter les attaques DoS et DDoS, tout en offrant une sensibilité accrue et une réduction des taux de fausses alarmes.

Mots-clés : TCP/IP, Cyberattaques, attaques DoS/DDoS, attaques TCP SYN, carte EWMA, carte EWMA adaptative, Base de données DARPA99.

General Introduction

General Introduction

In today's digitally intertwined landscape, networks serve as the lifeblood of modern society, facilitating seamless communication, data exchange, and business operations on a global scale. From personal interactions to enterprise transactions, the reliance on networks is pervasive, underpinning virtually every aspect of daily life. However, this interconnectedness also exposes networks to a multitude of threats, chief among them being cyber-attacks.

Denial of Service (DoS) and its more insidious variant, Distributed Denial of Service (DDoS), loom large as redoubtable challenges to network security. These attacks, orchestrated by malicious actors, involve overwhelming network resources with a flood of traffic, rendering services inaccessible to legitimate users. The repercussions of such attacks can be severe, encompassing financial losses, reputational damage, and operational disruptions for affected entities.

To mitigate the risks posed by these attacks, the development of robust detection mechanisms is imperative. Among the arsenal of defense strategies, control charts emerge as a promising toolset. Widely utilized in diverse industries to monitor process variations and deviations, control charts offer a systematic approach to anomaly detection and trend analysis.

In cyber-security, researchers are extensively studying control charts to adapt their characteristics and enhance overall detection capabilities. The designed approaches initially indicated heightened sensitivity, reduced false positives, and improved resilience against evolving attack methodologies.

In this study, we embark on an exploration of the effectiveness of some models of Adaptive EWMA (Exponentially Weighted Moving Average) control chart compared to the conventional EWMA chart in detecting DoS and DDoS attacks within network traffic. Our research methodology entails a comprehensive examination of network architectures, cyber-attack methodologies, and control chart principles. By leveraging datasets such as DARPA99, we aim to simulate various attack scenarios and evaluate the performance of Adaptive control chart models alongside traditional EWMA chart.

The methodology of this dissertation involves a comprehensive study of networks, cyber-attacks, and control charts. The research will be divided into three chapters as follow:

The first chapter delves into the fundamental aspects of networks and cyber-threats. It begins by defining networks and exploring the TCP/IP network architecture, along with essential protocols like TCP, UDP, and ICMP. The chapter then delves into cyber-security, covering vulnerability types and various cyber-attacks such as phishing, malware, DoS, DDoS, brute force, and social engineering attacks. Notably, it examines specific DoS and DDoS attack types like TCP SYN flood, UDP flood and Smurf. Finally, the chapter introduces common protection and detection methods essential for enabling network defenses against evolving threats.

The second chapter is about control charts and particularly some models of Adaptive EWMA charts. It covers the history of control charts, their definition, applications and description of traditional control charts such as Shewhart, CUSUM and EWMA. Then it focuses on different variant of EWMA chart, namely D-EWMA (Double EWMA), T-EWMA (Triple EWMA), FSS-EWMA (Fixed Sample Size EWMA), VSS-EWMA (Variable Sample Size EWMA), VSSILF-EWMA (Variable Sample Size EWMA as an Integer Linear Function) and adaptive ASW-EWMA (Adaptive Sliding Window EWMA).

In the third chapter, we provide a comprehensive explanation of the procedure employed for utilizing EWMA chart and its Adaptive models in detecting DoS and DDoS attacks. An assessment of their performance under various types of attacks; culminating in a comparative study conducted among these charts.

Chapter I

Cyber-attacks in IP Networks

Chapter I Cyber-attacks in IP Networks

I.1. Introduction

The emergence of the Internet has led to a profound transformation in communication patterns, information exchange, and global-scale information storage, given its usage in all fields. However, this digital revolution inevitably brings forth new challenges, including the rapid growth of cyber-attacks, specifically targeting Internet network. This rapid development highlights the necessity to understand and confront the threats to the security and stability of these communication networks, which are essential to our modern era.

This chapter provides an introduction to networks, their architectures and protocols, cybersecurity, types of cyber-attacks, with a particular focus on detecting DoS and DDoS attacks.

I.2. IP Networks

An IP (Internet Protocol) network is a global interconnected system of computing devices using standard communications protocols, such as TCP/IP. It allows sharing of information, access to online services, and communication between users located anywhere in the world. This network is based on a varied infrastructure including submarine cables, satellites, wireless networks, and data centers.

An IP address, or Internet Protocol address, is a numeric series assigned to each device connected to a computer network using the Internet Protocol for communication. This address plays a crucial role in routing data through the network. There are two main versions of IP addresses, namely IPv4 (Internet Protocol version 4) and IPv6 (Internet Protocol version 6) **[1] [2]**.

I.3. TCP/IP definition

For two machines to share data within a network, they must follow a defined set of regulations governing all facets of data transmission. These regulations, known as protocols, dictate the parameters of communication **[3]**.

Officially named the TCP/IP Internet Protocol Suite and commonly referred to as TCP/IP, it’s a suite of communication protocols that can be used to communicate across any set of interconnected networks. At its core, TCP/IP consists of two main protocols: Transmission Control Protocol (TCP) and Internet Protocol (IP) **[4]**.

Transmission Control Protocol (TCP): TCP handles the reliable transmission of data by breaking it into packets, ensuring that they reach their destination accurately and in the correct order.

Internet Protocol (IP): is responsible for addressing and routing packets across networks. It assigns unique IP addresses to devices and determines how data packets should be forwarded from the source to the destination across interconnected networks **[5]**.

I.4. TCP/IP Network Architecture

The TCP/IP protocol stack represents a structured arrangement of protocol layers designed for networks and systems.

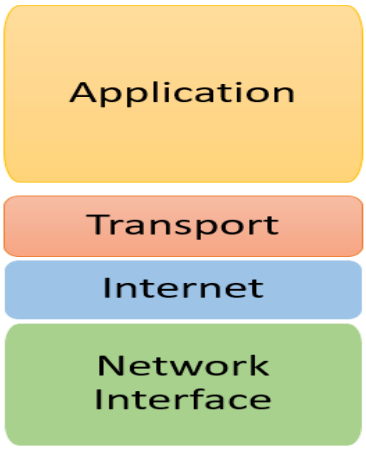


Figure I.1: The 4 layers of TCP/IP Architecture **[3]**

- **The network access layer:** this layer handles the transmission of data among devices within a local network. It facilitates data transmission over physical media like cable or Wi-Fi by employing protocols such as Ethernet and Wi-Fi.
- **The internet (network) layer:** This layer manages the routing of data packets between different networks. It utilizes IP addresses to identify hosts and networks, employing routing protocols like OSPF (Open Shortest Path First) and RIP (Routing Information Protocol).
- **The transport layer:** This layer ensures end-to-end data transmission, providing mechanisms for flow control and error correction to ensure reliable communication. Commonly used protocols at this layer include TCP (Transmission Control Protocol) and UDP (User Datagram Protocol)
- **The application layer:** This layer oversees high-level services like email, file transfer, and web browsing. It employs protocols such as HTTP (Hyper Text Transfer Protocol), FTP (File Transfer Protocol), and SMTP (Simple Mail Transfer Protocol).

I.5. Protocols used in TCP/IP Networks

I.5.1. UDP (User Datagram Protocol)

In the TCP/IP protocol suite, the User Datagram Protocol or UDP provides the primary mechanism that application programs use to send datagrams to other application programs **[4]**.

UDP offers protocol ports that aid in distinguishing among various programs running on a single machine. Each UDP message includes both a destination port number and a source port number, facilitating accurate delivery to the intended recipient at the destination and enabling the recipient to respond accordingly.

The UDP provides an unreliable delivery service using IP to transport messages between machines. An application program that uses UDP accepts full responsibility for handling the problem of reliability, including message loss, duplication, delay **[6]**.

Each UDP message is called a user datagram. Conceptually, a user datagram consists of two parts: a UDP header and a UDP data (Datagram) area.

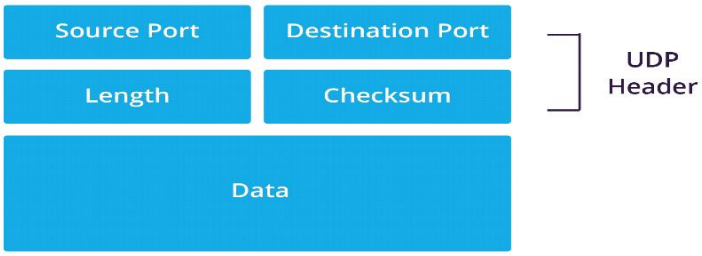


Figure I.2: UDP datagram format

the header is divided into four 16-bit fields that specify the port from which the message was sent, the port to which the message is destined, the message length, and a UDP checksum.

The Source Port and Destination Port fields contain 16-bit UDP protocol port numbers. These numbers are utilized to demultiplex datagrams among the processes that are awaiting reception.

The Length field indicates the number of octets in the UDP datagram, Consequently, the minimum value for Length is eight, representing the length of the header alone.

The UDP checksum is not mandatory and can be entirely omitted; a value of zero in the Checksum field indicates that the checksum has not been calculated **[6]**.

I.5.2. TCP (Transmission Control Protocol)

Transmission Control Protocol is one of the main protocols of the TCP/IP suite and is responsible for establishing and maintaining connections between devices on the networks.

TCP ensures reliable, ordered, and error-checked delivery of data packets over IP networks. It also provides mechanisms for establishing connections, breaking data into packets, sequencing packets, acknowledging receipt of packets, and retransmitting lost packets, making it a fundamental protocol for reliable communication over the internet **[7]**. Figure I.3 gives the TCP's three-way handshake process for enabling a normal client/server connection

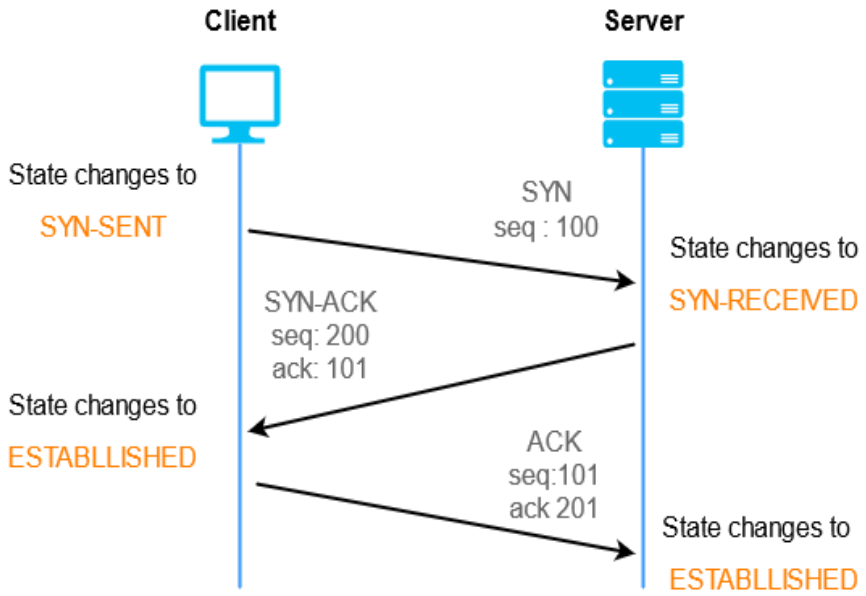


Figure I.3: TCP's three-way handshake

- **SYN**: the client initiates a connection request to the server by transmitting a segment containing the SYN (Synchronize Sequence Number) flag. This flag notifies the server of the client's intention to commence communication and indicates the sequence number from which it will begin sending segments.
- **SYN + ACK:** the server acknowledges the client's request by sending a segment with the SYN-ACK signal bits set. The acknowledgment (ACK) confirms receipt of the segment, while SYN indicates the sequence number from which the server intends to initiate segments.
- **ACK:** the client acknowledges the server's response finalizing the establishment of a reliable connection. This connection serves as the foundation for initiating the actual data transfer between the client and the server.

I.5.3. ICMP (Internet Control Message Protocol)

ICMP protocol serves as an error reporting mechanism, enabling routers that encounter errors to communicate them back to the original source.

When a datagram causes an error, ICMP can only report the error condition back to the original source of the datagram; the source must relate the error to an individual application program or take other action to correct the problem **[4]**. The ICMP header is organized as follow:

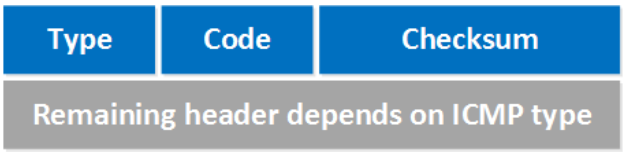


Figure I.4: ICMP messages format

The first byte specifies the type of ICMP message. For example, type 8 is used for an ICMP echo request and type 0 is used for an ICMP echo reply.

The second byte called code specifies what kind of ICMP message it is.

The third field are 2 bytes that are used for the checksum to see if the ICMP header is corrupt or not.

The table below shows some ICMP Messages

Type	Code	Message
0	**0**	**Echo Reply (ping)**
3	**2**	**Destination Protocol unreachable**
3	**7**	**Destination host unknown**
8	**0**	**Echo Request (ping)**
10	**0**	**Router Discovery**
12	**0**	**Bad IP header**

Table I.1: ICMP messages types

I.6. IP networks security

Security in an Internet network refers to the set of measures and protocols implemented to protect data, systems, and communications against potential threats, including cyber-attacks, intrusions, and loss of confidentiality. It encompasses the implementation of strategies such as data encryption, user authentication, access management, firewalls, and intrusion detection systems to ensure the integrity, availability, and confidentiality of information circulating on the network. The importance of security in an Internet network is

crucial as global connectivity exposes systems to considerable risks. Cyber-attacks are becoming increasingly sophisticated, making security a priority to ensure user trust, protect privacy, and maintain operational continuity **[8] [9]**.

I.6.1. Cyber-security definition

Cyber-security encompasses a comprehensive set of practices, tools, and methodologies aimed at safeguarding digital systems and assets from malicious activities and unauthorized access. It entails the deployment of proactive measures such as encryption, access controls, and threat detection mechanisms to uphold the confidentiality, integrity, and availability of sensitive information within digital environments. As a dynamic and evolving discipline, cyber-security addresses an array of cyber threats, including but not limited to cyber-attacks, data breaches, and malware infections, with the overarching goal of fortifying the resilience and security posture of digital infrastructures **[10]**.

I.6.2. The objectives of cyber-security

The objectives of cyber-security encompass a comprehensive set of goals aimed at protecting digital assets and ensuring the secure operation of information systems. These objectives collectively contribute to building a resilient defense against cyber threats, covering confidentiality, integrity, availability, authentication, non-repudiation, resilience and compliance **[11]**.

1. Confidentiality: Confidentiality ensures that sensitive information remains accessible only to authorized individuals or systems. Encryption, access controls, and secure communication protocols are employed to prevent unauthorized access.

2. Integrity: Integrity focuses on maintaining the accuracy and trustworthiness of data and systems. It involves preventing unauthorized alterations, ensuring data consistency, and maintaining the reliability of critical systems.

3. Availability: Availability ensures that information and services are accessible when needed. Cyber-security measures aim to prevent disruptions caused by cyber-attacks, system failures, or other incidents.

4. Authentication: Authentication verifies the identity of users, devices, or systems to prevent unauthorized access. Strong authentication mechanisms, including multi-factor authentication, play a vital role in cyber-security.

5. Non-repudiation: Non-repudiation ensures that the origin or delivery of information cannot be denied by the involved parties. Digital signatures and audit trails are common tools to achieve non-repudiation.

6. Resilience: resilience is to allow computer systems and networks to continue functioning despite failures, errors, or attacks. Security measures for resilience may include redundancy systems, business continuity plans, and activity recovery plans to minimize the impact of service interruptions.

7. Compliance: compliance is to ensure that computer systems, networks, data, and information comply with current laws and regulations. Security measures for compliance may include the implementation of security policies, compliance control procedures, and user training to ensure that rules and standards are followed.

Understanding and implementing these cyber-security objectives is essential for organizations and individuals to navigate the complex landscape of cyber threats. The suggested resources provide in-depth insights into cryptography, ethical hacking, cyber resilience, authentication, and security engineering, offering a comprehensive understanding of cyber-security principles and practices **[12]**.

I.7. Network Vulnerabilities

1.7.1. Definition

Vulnerabilities are weaknesses, gaps, or flaws in computer systems, applications, networks, or configurations that can be exploited by attackers. They represent potential entry points for security compromises, whether caused by programming errors, design flaws, improper configurations, or other factors **[13]**.

I.7.2. Types of vulnerabilities

There are various types of vulnerabilities grouped into three main categories **[14]**:

1. **Physical vulnerabilities:** These vulnerabilities are related to physical security issues that enable an attacker to physically access a system or organization. For example, they may include theft of computer equipment, bypassing locks, or unauthorized access to protected areas.

2. **Organizational vulnerabilities:** These weaknesses occur when an organization's policies, procedures, or practices are not robust enough to protect its information technology systems against potential threats. Examples include insufficient user access management, lack of security training, or absence of contingency plans.

3. **Technological vulnerabilities:** These weaknesses are present in computer systems, applications, and networks. Examples include programming errors, security issues, incorrect configurations, or coding errors.

I.8. Cyber-attacks

Cyber-attacks are deliberate and malicious actions taken by individuals, groups or entities, aimed at exploiting computer vulnerabilities to compromise the confidentiality, integrity or availability of data, computer systems or networks. These attacks can take various forms, such as denial of service attacks, phishing, malware, and are often motivated by financial, political, ideological, or other malicious motives **[15]**.

I.8.1. Types of Cyber-attacks

I.8.1.1. DoS and DDoS Attacks

Denial of service (DoS) attacks and distributed denial of service (DDoS) attacks represent serious threats to the availability of online services. DoS attacks are malicious attempts to render a service, system, or network unavailable by deliberately overwhelming its resources, usually through a single attack source. In contrast, DDoS attacks involve a multitude of sources, often geographically distributed, cooperating to flood the target with traffic, exceeding its normal capacity. Attackers exploit vulnerabilities in network protocols, web services, or applications to saturate resources and make the service inaccessible to legitimate users. Detecting these attacks requires advanced traffic monitoring, behavioral analysis, and anomaly detection tools to spot unusual patterns. DDoS protection solutions, such as traffic filtering, load balancing, and the use of specialized services, are essential methods to effectively mitigate these attacks in real time **[16].**

I.8.1.2. Phishing

Phishing attacks exploit user trust by impersonating trusted entities. Cyber-criminals create persuasive messages, often in the form of deceptive emails, impersonating legitimate institutions. These messages mislead recipients into revealing sensitive information such as

login credentials or credit card details. Typically, rush tactics, fake website creation, and emotional manipulation are used to quickly trick victims into providing confidential data **[19]**.

Vigilance and awareness are key to detecting phishing. Users should carefully check the details of the message, such as the URL of included links, the quality of spelling and grammar, and the authenticity of the sender. Technology solutions, like phishing filters and security software, can also help identify and block suspicious emails or websites **[17]**.

I.8.1.3. Malware

Malware attacks, such as viruses, worms and Trojan horses, aim to compromise the security of systems by exploiting vulnerabilities or infiltrating networks. Viruses spread by attaching themselves to executable files, worms reproduce themselves, and Trojans deceive users by pretending to be legitimate software. These threats can be introduced via infected email attachments, malicious downloads, or compromised websites. Once on the system, the malware can steal data, disrupt operations, or even allow attackers to access and control the system remotely **[19]**.

Early detection of malware attacks is crucial. Using up-to-date antivirus and anti-malware software is fundamental to scanning and identifying malware. Proactively monitoring system behavior, detecting network traffic anomalies, and implementing robust firewalls can also help spot suspicious activity. Users should be trained to recognize red flags, such as unusual system performance, suspicious pop-ups, or unexplained file changes **[18]**.

I.8.1.4. Social Engineering Attacks

Social engineering attacks aim to manipulate individuals into disclosing confidential information using persuasion tactics rather than technical vulnerabilities. Attackers exploit people's trust by using psychological manipulation. They can create false scenarios, feign legitimate authority, or even pose as trusted others. These attacks take various forms, such as sophisticated phishing emails, fraudulent phone calls, or even in-person interactions where attackers pose as co-workers or employees. The ultimate goal is to mislead victims into disclosing sensitive information like passwords, credit card numbers, or other confidential data **[8]**.

Detecting social engineering attacks relies on awareness and education. Users should be trained to spot warning signs such as unusual requests for confidential information, linguistic errors in communications, or emotional pressure to act quickly. Organizations can strengthen

security by implementing strict identity verification protocols, promoting a culture of security, and using security solutions that can detect social engineering attempts **[19]**.

I.8.1.5. Man-in-the-Middle (MITM) attacks

Man-in-the-middle attacks exploit weaknesses in communication between two legitimate parties. The attacker secretly inserts himself between the two parties, thus intercepting all data exchanged. This can happen at the network level, where an attacker can use techniques such as packet interception or ARP (Address Resolution Protocol) poisoning to direct traffic through it. Alternatively, MITM attacks can occur at the application layer, where techniques such as phishing or creating fake Wi-Fi hotspots are used to deceive users and redirect them to channels controlled by the attacker **[9]**.

Detecting MITM attacks requires proactive vigilance. Users and administrators should watch for signs of suspicious activity, such as unexpected changes in communications, warnings of invalid certificates, or inexplicable delays in data transmission. Using security protocols such as communications encryption (like HTTPS) can also help mitigate the risk of MITM by making it more difficult for an attacker to read or modify data **[8]**.

I.8.1.6. SQL Injection Attacks

SQL injection attacks exploit vulnerabilities in web applications by injecting malicious SQL commands into database queries. These attacks can occur when web applications do not properly validate input data or do not sanitize user-provided data. Attackers exploit this weakness by inserting unauthorized SQL code into input fields or URL parameters **[19]**.

The way the attack works is that web applications often process input data insecurely, allowing attackers to add additional SQL commands to those already present in the query. For example, an SQL attack may modify a data select query to retrieve sensitive information such as passwords, or it may alter an update query to modify database contents **[20]**.

To detect SQL injection attacks, developers can implement secure coding practices such as using prepared queries with bound parameters, employing user input validation mechanisms, and limiting permissions. Database access for accounts used by the application **[8]**.

Security tools such as web application firewalls (WAFs) can also help detect and block SQL injection attacks by monitoring incoming traffic and identifying suspicious patterns or attempts to inject malicious code.

I.8.1.7. IP Address Spoofing attacks

IP spoofing attacks involve forging the source IP address in data packets to mislead the recipient about the identity of the sender. This technique allows an attacker to pass off the traffic they generate as coming from a legitimate source. IP spoofing attacks are often used in Man-in-the-Middle (MITM) attacks where the attacker can intercept, modify, or simply observe communications between two parties **[9]**.

The way the attack works relies on the fact that many communications protocols, such as Internet Protocol (IP), do not provide built-in mechanisms for verifying the authenticity of source IP addresses. The attacker can thus send packets with a falsified IP address, misleading the recipient about the real origin of the traffic.

Detecting IP spoofing attacks can be done using techniques such as validating IP headers, setting up IP address filters in firewalls to block suspicious IP addresses, or the use of advanced intrusion prevention technologies **[19]**.

I.8.1.8. Border Gateway Protocol (BGP) vulnerability Attacks

Border Gateway Protocol (BGP) vulnerability attacks aim to manipulate routing information on the Internet by forging BGP announcements. BGP is the protocol used to exchange routing information between autonomous systems (AS) on the Internet. BGP attacks typically involve broadcasting fake BGP announcements that declare the availability of routes to certain destinations. This can result in legitimate traffic being redirected to malicious destinations controlled by the attacker **[21]**.

The attack's operation relies on the intrinsic trust of the BGP protocol, which assumes that routing announcements come from trusted sources. Attackers exploit weaknesses in BGP by serving fake announcements, tricking routers on the network into redirecting traffic to malicious paths.

To detect BGP attacks, various techniques are used. Monitoring BGP traffic, analyzing sudden changes in routing tables, and authenticating BGP announcements by security mechanisms such as Resource Public Key Infrastructure (RPKI) can help identify anomalies and manipulation attempts.

I.8.1.9. Brute force Attack

A brute force attack is a cyber-criminal method where an attacker attempts to guess a password or authentication key by systematically trying all possible combinations. This can be done manually or using automated programs that test thousands or even millions of combinations in a very short time. The goal is to compromise a user account, system, or network. To detect a brute force attack, administrators can monitor activity logs for repeated or suspicious attempts. Protecting against such attacks involves using strong passwords that are changed regularly and implementing security policies that limit the number of login attempts **[9]**.

I.8.2. The objectives of cyber-attacks

Cyber-attacks can have various objectives depending on the motivations of the attackers. Here are some of the main goals of cyberattacks **[22]**:

- **Theft of sensitive information:** Attackers often seek to access, copy, or steal confidential data, such as personal information, financial data, trade secrets, or intellectual property.
- **Extortion**: Some attackers use ransomware to encrypt an organization or individual's files, then demand payment of a ransom in exchange for the decryption key.
- **Sabotage:** Certain groups or individuals may carry out attacks to deliberately disrupt or damage systems, networks or critical infrastructure.
- **Espionage:** Governments, criminal organizations or activist groups may carry out cyber-attacks with the aim of spying on targeted entities and obtaining strategic information.
- **Industrial cyber espionage:** Attackers may aim to obtain information about a company's business practices, development plans, or trade secrets for competitive advantage.
- **Counter-espionage:** Some states or organizations may carry out cyber-attacks to counter the espionage activities of other entities.
- **Online activism:** Activist groups may carry out attacks to express political, social, or ideological demands, disrupting online services to draw attention to their causes.
- **Development of botnets:** Some attackers create botnets by infecting numerous devices to control them remotely, often with the aim of carrying out coordinated attacks.

- **Ideological purposes:** Certain individuals or groups may carry out attacks to promote their ideas, propagate political messages, or challenge political regimes.

I.9. Examples of DoS and DDoS attacks

I.9.1. TCP SYN flood attack

The TCP SYN flood attack is a type of DoS attack that exploits a vulnerability in the TCP communication protocol. When a client intends to establish a TCP connection with a server, it initiates the process by sending a SYN segment to the server. Upon receiving the SYN segment, the server responds with a SYN-ACK segment signals its readiness to establish the connection. Finally, the client sends an ACK (acknowledgment) segment to finalize the connection **[23]**.

During a TCP SYN flood attack, the attacker deliberately floods the target server with a large volume of SYN segments. Often, the attacker employs spoofed IP addresses to obfuscate their identity. Subsequently, the server responds to each SYN segment with a SYN-ACK segment. However, the attacker neglects to respond with the final ACK segment to complete the connection. Consequently, the server becomes inundated with numerous pending TCP connections, consuming its resources and impeding its ability to process legitimate requests.

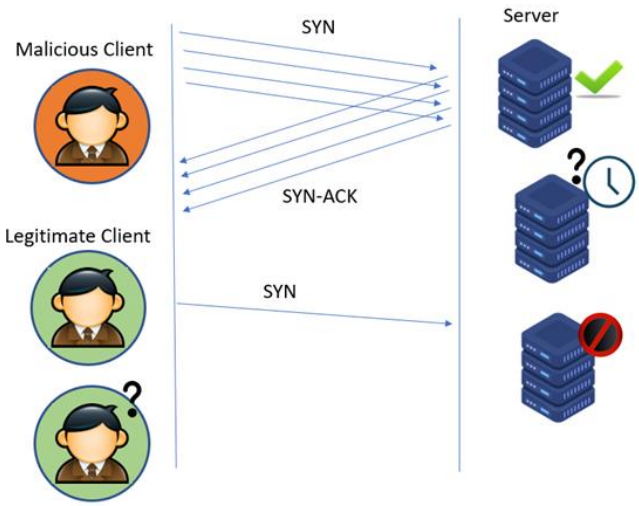


Figure I.5: DoS TCP SYN flood attack

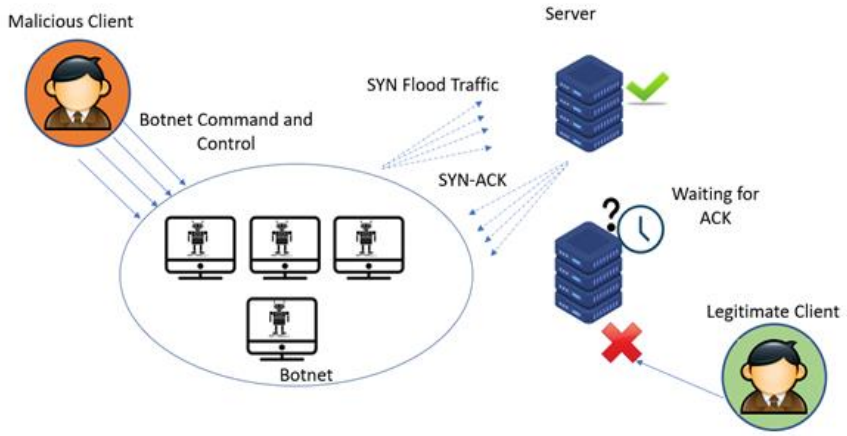


Figure I.6: DDoS TCP SYN flood attack

I.9.2. UDP flood attack

The UDP flood attack, a variant of the distributed denial of service (DDoS) attack, utilizes the User Datagram Protocol (UDP) to inundate a target with network traffic. Unlike the TCP SYN Flood attack that targets the Transmission Control Protocol (TCP), the UDP Flood attack aims to saturate the target network's bandwidth by flooding it with a high volume of UDP packets.

In a UDP Flood attack, the attacker sends a multitude of UDP packets to the target using spoofed IP addresses. These UDP packets are directed to random ports on the target, leading to the exhaustion of the target's network resources. Since UDP does not require a connection establishment process before data transmission, the packets can overwhelm the target without any need for acknowledgment, causing a disruption in service for legitimate users **[23]**.

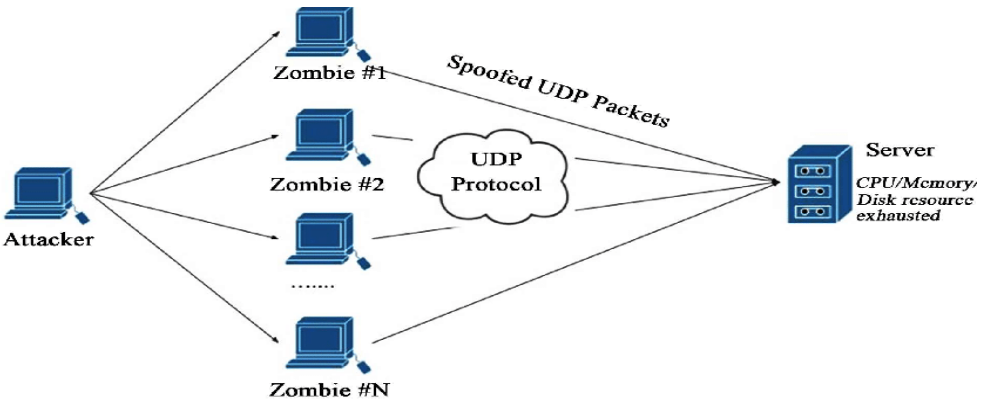


Figure I.7: DDoS UDP flood attack

I.9.3. Smurf Attack

The Smurf attack is a form of Distributed Denial of Service (DDoS) attack that exploits the characteristics of the ICMP protocol to flood a target with ping requests (ICMP Echo Request) using an amplification technique. This attack leverages the nature of ICMP, which allows a single request to trigger a response to multiple hosts, thus creating an amplification of the attack.

In a Smurf attack, the attacker sends ICMP ping packets with the victim's falsified IP address to a broadcast or multicast network, such as an open local network. Each host on this network, believing the ping request is from the victim, responds with an ICMP Echo Reply packet. When numerous hosts respond simultaneously, it can lead to saturation of the victim's bandwidth and eventually cause a service disruption **[23]**.

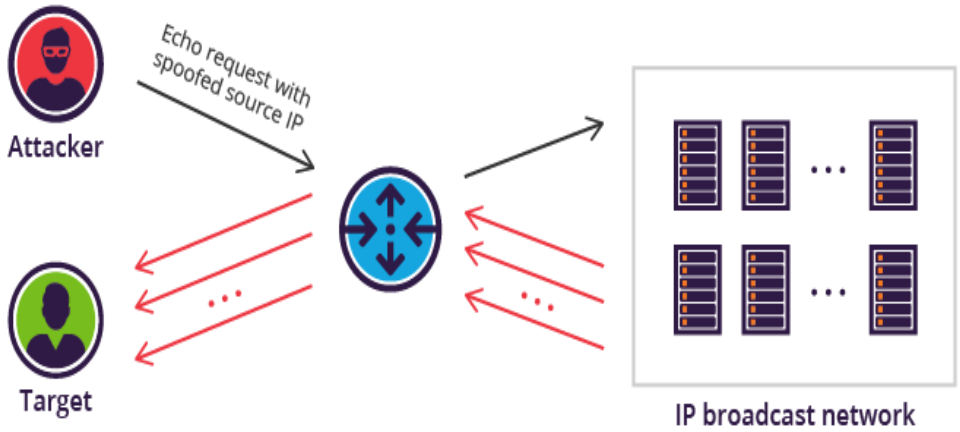


Figure I.8: DDoS Smurf attack

I.10. Conclusion

In this chapter we covered the critical aspects of networks and cyber-security, focusing on their pivotal role in safeguarding against cyber threats. Understanding the architecture and protocols of TCP/IP networks is highlighted as essential for identifying vulnerabilities and bolstering defenses against various attack vectors. The chapter explores the objectives and methods of cyber-attacks, including phishing, malware, brute force, social engineering, and DoS and DDoS attacks such as TCP SYN flood, UDP flood and Smurf. The devastating impact of cyber assaults on networks, ranging from data loss to financial and reputational damage, underscores the urgency of effective detection mechanisms, particularly for DoS and

DDoS attacks. Proactive security measures like intrusion detection systems, firewalls, and continuous network monitoring are emphasized as crucial for countering evolving cyber threats and safeguarding IP networks effectively.

To protect against cyber-attacks, the researcher has studied and developed techniques using control charts for detection and prediction. Adaptive models have been created to improved effectiveness. The following chapter explains traditional control charts and introduces some adaptive models based on EWMA charts.

Chapter II

Control Charts

Chapter II Control Charts

II.1. Introduction

In cyber-security, control charts constitute an interesting tool for revealing and mitigating cyber-attacks like DoS and DDoS. They meticulously monitor network traffic patterns, swiftly identifying deviations indicative of malicious activities. Among these, the Exponentially Weighted Moving Average (EWMA) chart stands out for its sensitivity to subtle shifts while minimizing noise. As threats evolve, the emergence of adaptive EWMA charts heralds a proactive approach, dynamically adjusting to emerging patterns for enhanced defense.

This chapter presents general concepts of control charts and focus with more details on EWMA-based adaptive models which are designed to deal with the common limitations of conventional EWMA chart.

II.2. Control chart history

Walter Shewhart is widely recognized as the father of modern statistical process control. In the 1920s, he developed the first control chart, which he used to monitor and improve the quality of telephone equipment produced by the Western Electric Company **[24]**. Shewhart's work revolutionized quality control and laid the foundation for modern statistical process control. He also developed the concept of common and special causes of variation, which is still used in statistical process control today **[25]**.

Over the years, control chart techniques have continued to evolve and improve. In the 1950s, several improvements were made to the basic control chart, including the development of the cumulative sum (CUSUM) chart and the exponentially weighted moving average (EWMA) chart. These charts were designed to detect small or gradual changes in a process. In the 1980s, multivariate control charts were developed to monitor processes with multiple variables. In recent years, with the rise of big data and machine learning, new techniques for control chart analysis have been developed, including artificial neural networks and fuzzy logic **[26]**.

During World War II, control charts gained widespread acceptance and played a pivotal role in defense manufacturing, ensuring the consistent quality of military equipment and supplies. This period of industrial mobilization saw the extensive application of statistical

methods, including control charts, to meet the stringent demands of wartime production **[27] [28]**. The post-war era witnessed further development and popularization of control chart techniques by statisticians such as W. Edwards Deming and Joseph M. Juran. Deming's seminal works, including "Out of the Crisis" and "Quality, Productivity, and Competitive Position," emphasized the importance of statistical process control as a key component of Total Quality Management (TQM). Juran, in his influential book "Juran's Quality Control Handbook," provided comprehensive insights into the principles and applications of control charts, further solidifying their status as indispensable tools in quality management **[29]**.

Today, control charts are widely used in many industries, including manufacturing, healthcare, finance, and telecommunications. They are a key tool in statistical process control, helping organizations to monitor and improve the quality of their products and services. With the increasing availability of data and the rise of data analytics, control charts are becoming more sophisticated and powerful. They are also being used in new and innovative ways, such as in the detection of cyberattacks and fraud **[25]**.

II.3. Control chart definition

Control charts are statistical tools employed to monitor and manage the stability of processes over time. These charts are crucial components of quality management systems, aiding in the detection of variations and deviations that may lead to defects in the final product or service. A control chart consists of three main elements: the data points, the center line, and the control limits. The data points are measurements of the process being monitored, while the center line represents the mean or average of these data points. The control limits, typically consisting of upper and lower bounds alongside the central line, provide a reference framework for analyzing the process's performance **[30]**.

Control charts follow a structured approach, beginning with the collection of relevant data for the observed process. Subsequently, control limits are established based on the inherent variability in the collected data, delineating the acceptable range within which the process should ideally operate. Data points are plotted on the control chart over successive time intervals, allowing practitioners to visually identify patterns or trends indicative of process stability or deviation **[25]**.

There are various types of control charts, such as X-bar (Average Chart) and R charts (range chart) for monitoring central tendency and variability, and attribute control charts for

assessing nonconformities. By incorporating control charts into quality management practices, organizations can foster a culture of continuous improvement, thereby driving long-term success and customer satisfaction. If a data point falls outside the predetermined control limits, it may indicate that the process is out of control and requires corrective action. This systematic approach to process monitoring and improvement is instrumental in enhancing product quality and operational efficiency **[31]**.

II.4. Applications of control charts

Control charts serve as vital tools across industries, aiding in monitoring processes and identifying variations that can impact business outcomes. In manufacturing and other sectors where precision is crucial, control charts swiftly detect deviations, enabling engineers to rectify issues promptly and maintain product quality. They can also be utilized beyond production lines, such as in monitoring employee satisfaction and budget adherence **[32]**.

In quality management, control charts play a pivotal role in ensuring processes meet standards by detecting both common and special cause variations. By promptly identifying deviations, organizations can take corrective actions, leading to consistent quality outputs and reduced scrap rates **[33].**

In healthcare, where quality is paramount, control charts aid in monitoring performance metrics like mortality rates and infection numbers. Techniques like Statistical Process Control (SPC) help healthcare providers detect anomalies early, allowing for timely interventions and improved patient outcomes **[34]**.

Moreover, control charts extend into personal life, offering a structured approach to managing goals, time, finances, and health metrics. Individuals can employ control charts to track progress, manage expenses, and monitor health parameters like blood pressure and cholesterol levels. Thus, control charts find application beyond industries, contributing to personal well-being and productivity **[32]**.

II.5. Shewhart Chart

The Shewhart control chart, named after its inventor Walter A. Shewhart, stands as a foundational tool in SPC. It serves to monitor the stability and performance of a process over time, providing a graphical representation of process data collected at regular intervals. This chart comprises key elements: a central line representing the process mean, and upper and lower control limits, which delineate the acceptable range of variation in the process. These

control limits are determined using specific equations derived from the process mean and standard deviation **[35] [36]**.

The equations for the control limits are as follows:

$$CL = \mu_0 \quad \textbf{(II.1)}$$

$$UCL = \mu_0 + 3\sigma \quad \textbf{(II.2)}$$

$$LCL = \mu_0 - 3\sigma \quad \textbf{(II.3)}$$

Here, CL denotes the central line, reflecting the process mean (μ_0),

$$\mu_0 = \frac{\sum_{i=1}^{n} Xi}{N} \quad \textbf{(II.4)}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (Xi - \mu_0)^2}{N}} \quad \textbf{(II.5)}$$

With: X_i is the data simple; N is the sample's size.

While UCL and LCL represent the upper and lower control limits, respectively. The upper and lower control limits are calculated by adding and subtracting three times the standard deviation (σ) from the process mean.

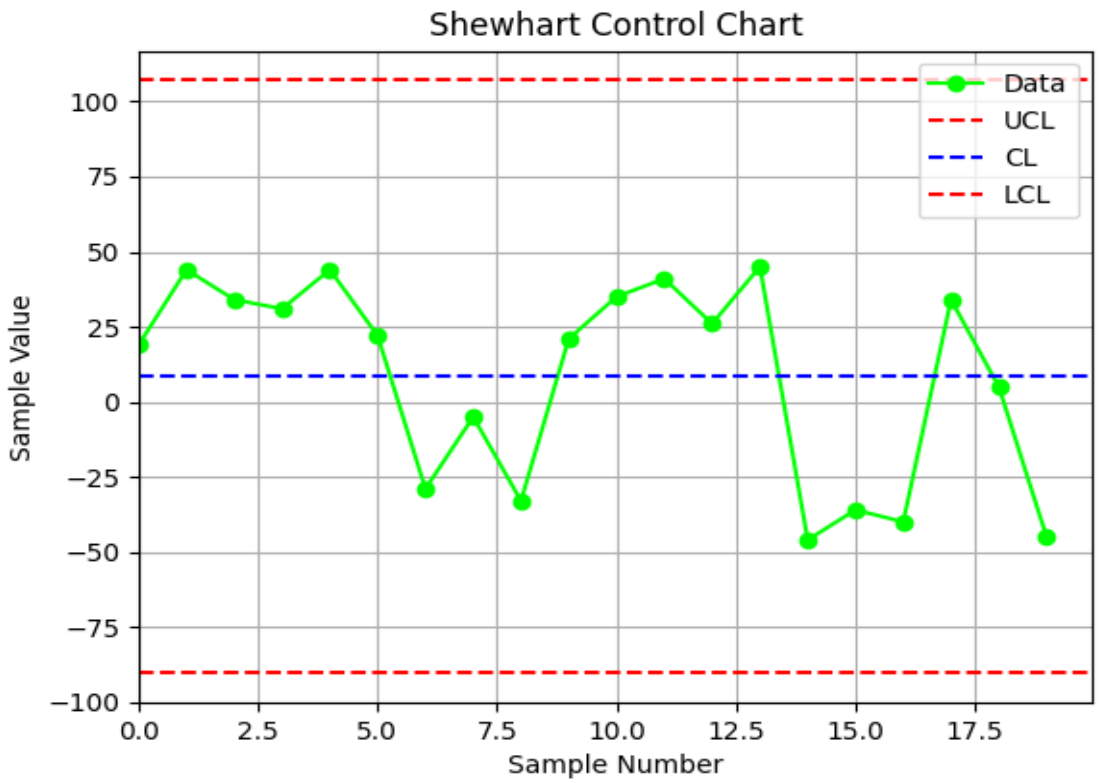


Figure II.1: Shewhart control chart for random data

In practical terms, when data points fall within these control limits, it indicates that the process is operating within expected variability, and any observed variation is likely due to common causes inherent in the process. However, if data points extend beyond these control limits, it suggests the presence of special causes of variation, which require further investigation and corrective action to restore process stability **[36]**.

While the Shewhart control chart is widely used for monitoring and controlling process quality in various industries, it does have limitations and assumptions that must be considered. One of the main assumptions is that the process data is normally distributed. If the data is not normally distributed, the chart may not be effective in detecting process changes. Another limitation is that the chart assumes that the process is in a state of statistical control, meaning that the process is stable and predictable over time. If the process is not stable, the chart may not be effective in detecting process changes. Despite these limitations, by utilizing Shewhart control charts, practitioners can systematically monitor processes, distinguish between common and special causes of variation, and take timely corrective measures as needed. This approach fosters continuous improvement and ensures consistent quality in

manufacturing, healthcare, finance, and other industries where process control is paramount **[37] [38]**.

II.6. CUSUM control chart

The Cumulative Sum (CUSUM) control chart, introduced by E.S. Page in 1961, stands as a cornerstone in SPC, offering a dynamic method for monitoring process performance. It excels in detecting subtle shifts in process mean or variance, making it adept at identifying nuanced changes often overlooked by traditional control charts. Much like the Shewhart control chart, the CUSUM chart is constructed using process data collected at regular intervals. However, instead of plotting individual data points, it calculates cumulative sums of deviations from the target value or process mean. This cumulative approach allows for more effective detection of performance shifts as deviations accumulate over time **[35] [36]**.

The CUSUM chart comprises two primary components: the cumulative sum of deviations (CUSUM) and decision intervals. Sequentially updated CUSUM values provide a running total of deviations from the target value, while decision intervals, determined based on predefined criteria, signal significant shifts in the process. Practically, the CUSUM chart offers early detection of process shifts by signaling when the cumulative sum exceeds a specified threshold. This proactive approach enables timely corrective action before significant deviations from the target occur. Additionally, the CUSUM chart exhibits greater sensitivity to small changes in process performance compared to traditional control charts **[35] [36]**.

The equations for the CUSUM chart are as follows:

The cumulative sum Ci up to the i-th sample is given by:

$$\boldsymbol{C_i = \sum_{j=1}^{i}(x_j - \mu_0)} \quad \textbf{(II.6)}$$

CUSUM defines two unilateral accumulations: positive C_i^+ and negative C_i^-:

$$\boldsymbol{C_i^+ = max\,[0, X_i - (\mu_0 + K) + C_{i-1}^+]} \quad \textbf{(II.7)}$$

$$\boldsymbol{C_i^- = max\,[0, (\mu_0 - K) - X_i + C_{i-1}^-]} \quad \textbf{(II.8)}$$

Where:

C_i^+: represents the accumulation of deviations above the target mean.

C_i^-: represents the accumulation of deviations below the target mean.

The initial values are: $C_0^+ = C_0^- = 0$.

K: is the sensitivity adjustment parameter of CUSUM, it chosen to limit the number of false alarms and is determined based on the significance of the gap that needs to be detected. A smaller value of "k" allows the method to detect minor changes, but also increases the chances of false alarms. Typically, the value of "k" is calculated using:

$$\boldsymbol{K = \frac{\sigma}{2}} \quad \textbf{(II.9)}$$

The control limits of the CUSUM are defined with the decision interval H:

$$\boldsymbol{H = 5\sigma} \quad \textbf{(II.10)}$$

When either of the two entities $Ci+$ or $Ci-$ exceeds H, the process is considered out of control.

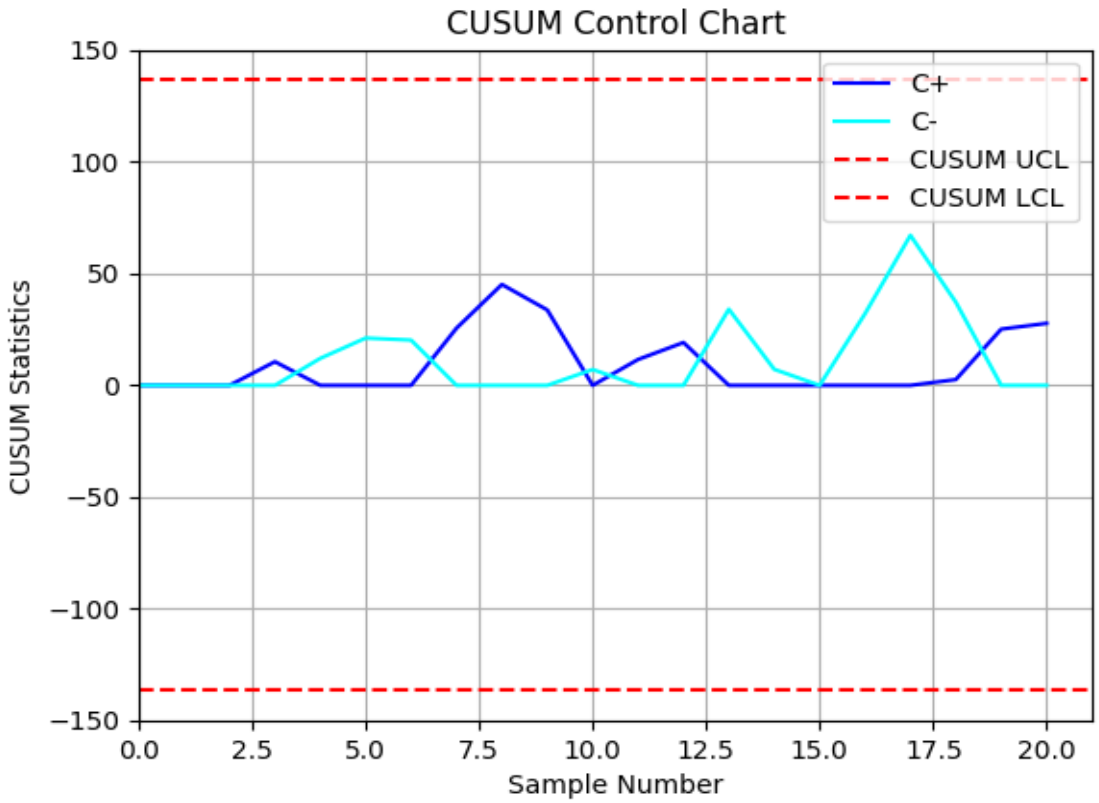


Figure II.2: CUSUM control chart for random data

Nevertheless, the CUSUM chart operates under certain assumptions and limitations. Like other control charts, it assumes known process data distribution and process stability

over time. Deviations from these assumptions can compromise the chart's effectiveness in detecting process shifts. However, incorporating CUSUM control charts into quality management practices enhances organizations' ability to monitor and improve process performance. This systematic approach fosters continuous improvement and ensures consistent quality outcomes across various industries **[37] [38]**.

II.7. EWMA Control Chart

Exponentially Weighted Moving Average (EWMA) control charts, pioneered by Roberts (1959), represent a pivotal advancement in SPC, providing a dynamic method for monitoring process stability over time. Originally developed as an enhancement to traditional Shewhart control charts, EWMA charts excel in detecting subtle shifts in process mean or variance. The underlying principle of EWMA charts lies in their weighting scheme, which assigns exponentially decreasing weights to historical data points while giving more weight to recent observations **[39] [40]**.

The construction of an EWMA chart involves calculating a weighted average of process measurements, incorporating both recent and historical data. The choice of the smoothing parameter (**λ**) is critical in EWMA chart construction, as it determines the rate of decay of weights. Smaller values of λ result in faster decay and greater responsiveness to recent changes. The EWMA statistic at time ***t*** is computed using the formula **[40]**:

$$\boldsymbol{Z_t = \lambda x_t + (1-\lambda) Z_{t-1}} \qquad \textbf{(II.11)}$$

where $\mathbf{Z_t}$ represents the EWMA statistic at time ***t***, $\boldsymbol{x_t}$ is the process measurement at time ***t***, and is $\boldsymbol{Z_{t-1}}$ the previous EWMA statistic.

Or by:

$$\boldsymbol{Z_i = \lambda x_i + (1-\lambda) Z_{i-1}} \qquad \textbf{(II.12)}$$

$$\boldsymbol{Z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^j Z_0} \qquad \textbf{(II.13)}$$

Where:

x_i: Value of the i^{th} sample.

Z_i: The EWMA of the sample.

Z_0: initial value of EWMA, typically chosen to be equal to $\boldsymbol{\mu_0}$ (the process mean).

The notation $\boldsymbol{Z_t}$ is commonly used when discussing the EWMA statistic in a time-series context. In this notation, ***t*** represents time, and $\boldsymbol{Z_t}$ denotes the EWMA value at a specific time point ***t***. This notation is often used when analyzing data over a period of time, such as in time-series analysis or when monitoring the performance of a process continuously over time.

On the other hand, the notation $\boldsymbol{Z_i}$ is typically used when discussing the EWMA statistic in the context of sample indices. Here, ***i*** represents the index of the sample, and $\boldsymbol{Z_i}$ represents the EWMA value corresponding to a specific sample ***i.*** This notation is more commonly used in statistical analyses involving discrete data points or samples, such as in quality control processes where samples are collected at distinct intervals.

In summary, both notations represent the EWMA statistic, but their usage depends on whether the analysis is focused on time-series data or discrete samples.

The control limits for EWMA charts are derived from the standard deviation of process data and are calculated using the following formulas **[41]**:

$$\boldsymbol{CL = \mu_0} \tag{II.14}$$

$$\boldsymbol{UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]}} \tag{II.15}$$

$$\boldsymbol{LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]}} \tag{II.16}$$

where $\boldsymbol{\mu_0}$ is the process mean, $\boldsymbol{\sigma}$ is the process standard deviation, ***L*** is a multiplier factor, typically set to 2.7 for control limits, and ***i*** represents the sample number.

λ: adjustment factor ranging between 0 and 1. It represents the weight assigned to different samples.

The closer λ is to 0: the weight $\lambda(1-\lambda)^{j}$ decreases slowly, and the past is taken into account more. This implies that small drifts will be more easily identified. However, abrupt drifts and significant disturbances will be less well detected.

The closer λ is to 1: $\lambda(1-\lambda)^{j}$ decreases rapidly, and the past is less taken into account. This implies that there will be a better reactivity to identify sudden disturbances but conversely, small variations will be less well detected.

If λ = 1: EWMA will be equivalent to the classical Shewhart control chart.

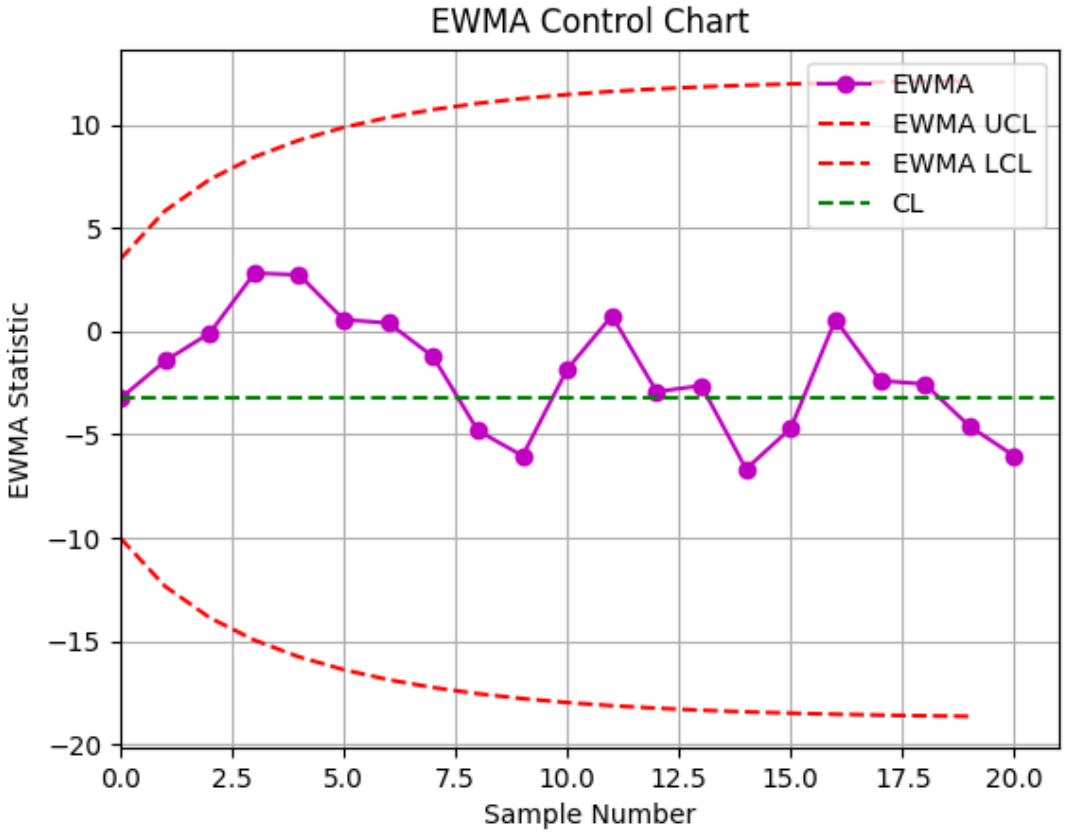


Figure II.3: EWMA control chart for random data

EWMA charts offer several advantages over traditional control charts, including improved sensitivity to small shifts and adaptability to various process characteristics. However, like any control chart method, EWMA charts also have limitations and assumptions that must be considered, such as the need for normally distributed process data and the assumption of process stability over time **[42]**.

Overall, EWMA control charts provide a balanced approach to process monitoring, offering both sensitivity to process shifts and stability in performance evaluation **[43]**. Their versatility and effectiveness make them widely used in industries such as manufacturing, healthcare, and finance, where continuous monitoring of process performance is essential for maintaining product quality and efficiency **[41] [14].**

II.8. Limitations of EWMA control Chart

The EWMA control chart is a useful tool for monitoring processes over time, but it does have certain limitations:

1. **Sensitivity to Parameter Choice:** The performance of the EWMA control chart is highly dependent on the choice of the smoothing parameter (often denoted as λ). Selecting an inappropriate value for λ can lead to either excessive sensitivity to small changes (resulting in false alarms) or insufficient sensitivity to significant shifts in the process (resulting in missed detections) **[25]**.

2. **Assumption of Normality:** Like many SPC methods, the EWMA control chart assumes that the data follow a normal distribution. If the underlying process does not meet this assumption (e.g., if it is highly skewed or has heavy tails), the control limits calculated based on normality may not accurately reflect the process variability, leading to incorrect interpretations of the chart **[35]**.

3. **Difficulty in Interpreting Control Limits**: Unlike traditional Shewhart control charts, which have fixed control limits based on the process mean and standard deviation, the control limits in an EWMA chart are not as straightforward to interpret. The control limits in an EWMA chart are based on a combination of the process mean, standard deviation, and the smoothing parameter, making their interpretation more complex **[35]**.

4. **Memoryless Property:** The EWMA chart assigns exponentially decreasing weights to past observations, which means that older observations have less influence on the current EWMA statistic. While this property allows the chart to be more responsive to recent changes in the process, it also means that the chart may not effectively capture longer-term trends or patterns in the data **[43]**.

5. **Limited Sensitivity to Small Shifts:** In some cases, the EWMA control chart may not be sensitive enough to detect small shifts in the process mean, especially if the process variability is high. This limitation can result in delayed detection of subtle changes, which may have practical implications for process improvement and quality control **[44]**.

II.9. D-EWMA (Double Exponentially Weighted Moving Average)

The D-EWMA control chart was developed to address certain limitations of the single EWMA control chart the aim is to provide enhanced sensitivity and responsiveness in detecting process changes, particularly in situations where autocorrelation is present. Autocorrelation refers to the correlation of a time series with its own past and future values, which is common in many industrial processes **[45]**.

The D-EWMA control chart is an extension of the EWMA control chart by performing exponential smoothing twice. Consider a sequence of independent and identically distributed random samples $\boldsymbol{X_1, X_2, \ldots, X_n}$, each drawn from a normal distribution with a mean $\boldsymbol{\mu_0}$ and a standard deviation σ. The D-EWMA control statistic $\boldsymbol{Z_i}$ is defined as **[45]**:

Firstly, we calculate the $\boldsymbol{Z_i}$ using the equation (**II.12**), then we use it to calculate $\boldsymbol{Y_i}$ as follow:

$$\boldsymbol{Y_i = \lambda Z_i + (1 - \lambda) Y_{i-1}} \qquad \textbf{(II.17)}$$

With $\boldsymbol{Z_i}$ and $\boldsymbol{Y_i}$, referring the EWMA statistic for the $\boldsymbol{i}^{\textbf{th}}$ sample, and $\boldsymbol{x_i}$ represents the measurement of the process for the $\boldsymbol{i}^{\textbf{th}}$ sample. $\boldsymbol{Z_{i-1}}$ and $\boldsymbol{Y_{i-1}}$ refers to the previous EWMA statistic, while $\boldsymbol{\lambda}$ serves as the smoothing parameter. Where $\boldsymbol{Y_0 = Z_0 = \mu_0}$ and smoothing constant $\boldsymbol{\lambda}$ is between **0** and **1.**

The average and variance of the DEWMA control chart are provided as follows:

$$\boldsymbol{\mu_Z = E(Z_i)} \qquad \textbf{(II.18)}$$

$$\boldsymbol{\sigma_z^2 = \frac{\lambda(2-2\lambda+\lambda^2)}{(2-\lambda)^3}\,\sigma_x^2} \qquad \textbf{(II.19)}$$

The upper and lower control limits are calculated using the following expressions:

$$\begin{cases} \boldsymbol{DUCL = \mu_Z + L\sigma_x\sqrt{\frac{\lambda(2-2\lambda+\lambda^2)}{(2-\lambda)^3}}} \\ \boldsymbol{DCL = \mu_Z} \\ \boldsymbol{DLCL = \mu_Z - L\sigma_x\sqrt{\frac{\lambda(2-2\lambda+\lambda^2)}{(2-\lambda)^3}}} \end{cases} \qquad \textbf{(II.20)}$$

When the D-EWMA statistic**,** $\boldsymbol{Y_i}$ ventures beyond the lower control limit (***LCL***) or upper control limit (***UCL***), indicating an out-of-control scenario, corrective action is initiated.

II.10. T-EWMA (Triple Exponentially Weighted Moving Average)

T-EWMA control chart is another tool used of SPC, with the sole purpose of monitoring and maintaining the stability of manufacturing and production processes.

Developed as a better version of the Double Exponentially Weighted Moving Average (D-EWMA) control chart by William H. Woodall, Kevin M. Grigg, and Sharon L. M. Poh in 1997.

D-EWMA which utilizes double exponential weighting, proved to be effective in detecting sudden shifts in process parameters. However, it struggled to discern gradual changes or trends often leading to delayed detection of process disturbances.to solve these shortcomings, the developers introduced T-EWMA as a more robust alternative **[46]**.

The T-EWMA design entails applying the smoothing procedure three times; instead of twice as for the D-EWMA or once as for the EWMA. T-EWMA control chart use the statistic **[47]**:

$$\begin{cases} \boldsymbol{Z_i = \lambda X_i + (1-\lambda) Z_{i-1}} \\ \boldsymbol{Y_i = \lambda Z_i + (1-\lambda) Y_{i-1}} \\ \boldsymbol{W_i = \lambda Y_i + (1-\lambda) W_{i-1}} \end{cases} \quad \textbf{(II.21)}$$

With $\boldsymbol{Z_i}$, $\boldsymbol{Y_i}$, $\boldsymbol{W_i}$, referring the EWMA statistic for the $\boldsymbol{i}^{\text{th}}$ sample, and $\boldsymbol{x_i}$ represents the measurement of the process for the $\boldsymbol{i}^{\text{th}}$ sample. $\boldsymbol{Z_{i-1}}$, $\boldsymbol{Y_{i-1}}, \boldsymbol{W_{i-1}}$ refers to the previous EWMA statistic, while $\boldsymbol{\lambda}$ serves as the smoothing parameter. Where $\boldsymbol{W_0 = Y_0}$, $\boldsymbol{Z_0 = \mu_0}$ and smoothing constant $\boldsymbol{\lambda}$ is between **0** and **1**.

The mean of the T-EWMA statistic is defined as:

$$\boldsymbol{E(W_i) = \mu_w} \quad \textbf{(II.22)}$$

For the large value of ***i,*** the variance of $\boldsymbol{Y_i}$ can be defined as:

$$\boldsymbol{\sigma_z^2 = \frac{6(1-\lambda)^6\lambda}{(2-\lambda)^5} + \frac{12(1-\lambda)^4\lambda^2}{(2-\lambda)^4} + \frac{7(1-\lambda)^2\lambda^3}{(2-\lambda)^3} + \frac{\lambda^4}{(2-\lambda)^2}} \quad \textbf{(II.23)}$$

The upper and lower control limits are calculated using the following expressions:

$$\begin{cases} TUCL = \mu_w + L\sigma\sqrt{\frac{6(1-\lambda)^6\lambda}{(2-\lambda)^5} + \frac{12(1-\lambda)^4\lambda^2}{(2-\lambda)^4} + \frac{7(1-\lambda)^2\lambda^3}{(2-\lambda)^3} + \frac{\lambda^4}{(2-\lambda)^2}} \\ TCL = \mu_w \\ TLCL = \mu_w - L\sigma\sqrt{\frac{6(1-\lambda)^6\lambda}{(2-\lambda)^5} + \frac{12(1-\lambda)^4\lambda^2}{(2-\lambda)^4} + \frac{7(1-\lambda)^2\lambda^3}{(2-\lambda)^3} + \frac{\lambda^4}{(2-\lambda)^2}} \end{cases} \quad \textbf{(II.24)}$$

Same as D-EWMA if Z_i set outs the lower and upper limits it signals the OOC (out of control) signal and calls the need to take action to correct the process.

II.11. Adaptive EWMA

Adaptive EWMA control charts represent a further development in statistical process control (SPC), offering a dynamic approach to monitoring process stability over time with the ability to adapt to changing process conditions. Initially introduced as an enhancement to traditional EWMA charts, Adaptive EWMA charts excel in detecting subtle shifts in process mean or variance, particularly in environments where process conditions are not constant **[35]**.

The Adaptive EWMA methodology maintains the underlying principles of EWMA charts, primarily focusing on the weighting scheme, which assigns exponentially decreasing weights to historical data while emphasizing recent observations. However, Adaptive EWMA introduces the capability to automatically adjust the smoothing parameter (**λ**) based on the current state of the process. This adaptation enables the chart to be more responsive to changes in process behavior, allowing for quicker detection of significant shifts or disturbances **[35][48]**.

The construction of an Adaptive EWMA chart involves calculating a weighted average of process measurements, similar to traditional EWMA. However, the choice of the smoothing parameter (**λ**) is now dynamic, adjusting in response to the observed process variability. Smaller values of **λ** result in faster decay and greater sensitivity to recent changes, while larger values provide more stability by giving less weight to recent observations **[35][48]**.

The Adaptive EWMA statistic at time t is computed using the formula in the equation (**II.11**):

Where:

- Z_t represents the Adaptive EWMA statistic at time ***t***.
- x_t is the process measurement at time ***t***.

- Z_{t-1} is the previous Adaptive EWMA statistic.
- λ is the adaptive smoothing parameter at time t.

The adaptive nature of the smoothing parameter allows the chart to adjust its sensitivity to changes in process conditions, balancing the trade-off between responsiveness and stability. This adaptability enhances the chart's effectiveness in various process environments, ensuring timely detection of deviations from the desired performance **[48]**.

The control limits for Adaptive EWMA charts are derived similarly to traditional EWMA charts, incorporating the process mean and standard deviation. However, the adaptive nature of the smoothing parameter influences the calculation of control limits, potentially leading to variations in their positions based on the current process conditions **[35]**.

Adaptive EWMA charts offer several advantages over traditional EWMA charts, including improved responsiveness to changing process conditions and enhanced sensitivity to subtle process shifts. However, like any control chart method, Adaptive EWMA charts have assumptions and limitations that must be considered, such as the need for normally distributed process data and the assumption of process stability over time **[35][48]**.

Overall, Adaptive EWMA control charts provide a powerful tool for process monitoring, offering a balance between sensitivity to process changes and stability in performance evaluation. Their adaptability and effectiveness make them valuable in industries where continuous monitoring of process performance is crucial for maintaining product quality and efficiency **[35][48]**.

II.12. FSS-EWMA (Fixed Sample Size EWMA)

Fixed Sample Size EWMA control charts stand as pillars in Statistical Process Control, tracing their origins to Roberts' seminal work in 1959. These charts serve as vigilant sentinels, guarding against subtle shifts in process parameters over time while offering a stable foundation for performance evaluation **[39]**.

At the heart of FSS-EWMA lies the computation of the EWMA statistic, a weighted average amalgamating the current process measurement with its historical counterparts. This statistic for the i^{th} sample is derived as the equation (**II.12**):

$\boldsymbol{Z_i}$ represents the EWMA statistic for the $\boldsymbol{i}^{\text{th}}$ sample, $\mathbf{x_i}$ denotes the process measurement for the $\boldsymbol{i}^{\text{th}}$ sample, $\boldsymbol{Z_{i-1}}$ is the previous EWMA statistic, and $\boldsymbol{\lambda}$ is the smoothing parameter or exponential weight constant **[49]**.

The determination of control limits is pivotal in delineating the boundaries of process stability. Derived from the process mean ($\boldsymbol{\mu_0}$) and standard deviation ($\boldsymbol{\sigma}$), these limits ensure sensitivity to process variations. For $\boldsymbol{i=0}$, $\boldsymbol{Z_i = Z_0}$ is starting value and is often taken equal to the process target value. The upper and lower control limits are calculated using the following expressions:

$$\begin{cases} \boldsymbol{UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}(1-(1-\lambda)^{2i})}} \\ \boldsymbol{CL = \mu_0} \\ \boldsymbol{LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}(1-(1-\lambda)^{2i})}} \end{cases} \quad \textbf{(II.25)}$$

Where, $\boldsymbol{n}$ is the number of samples in each subgroup and $\boldsymbol{L}$ is the coefficient of control limits which is chosen to obtain a desirable in control average run length (ARL_0). The EWMA control chart signals as soon as $\boldsymbol{Z_i} > UCL$ or $\boldsymbol{Z_i} < LCL$ **[49]**.

In the initial stage, the starting values for control limits are determined based on the process mean and standard deviation. In FSS-EWMA, adjustments are not made to these starting values, as the sample size remains fixed throughout the monitoring process **[39].**

The effectiveness of FSS-EWMA lies in its ability to provide timely detection of process shifts and disturbances while maintaining stability in performance evaluation. By utilizing a fixed sample size and control limits, FSS-EWMA charts offer robust monitoring capabilities, making them valuable tools in industries where continuous process surveillance is essential for quality assurance and efficiency **[39].**

II.13. VSS-EWMA (Variable Sample Size EWMA)

Variable Sample Size EWMA control charts represent a significant advancement in statistical process control, offering a dynamic approach to monitoring process stability while adapting to changing process conditions. These charts build upon the traditional EWMA methodology, integrating variable sample sizes to enhance sensitivity to small shifts in process parameters **[39]**.

The design of VSS-EWMA charts involves utilizing the same statistical formulation as the Fixed Sample Size EWMA (FSS-EWMA) charts, as mentioned in the previous subsection

[49]. However, in VSS-EWMA, additional considerations are made to adjust sample sizes dynamically based on the current process state. This adjustment aims to improve the chart's responsiveness to changes in process behavior, allowing for quicker detection of significant shifts or disturbances.

To implement VSS-EWMA, upper and lower warning limits (UWL and LWL) are introduced, along with the traditional control limits, to define safety and warning zones. These zones provide a visual representation of the process state relative to its control limits, aiding in decision-making during process monitoring **[49]**.

$$\begin{cases} \boldsymbol{UWL = \mu_0 + L_1\sigma\sqrt{\frac{\lambda}{n_i(2-\lambda)}(1-(1-\lambda)^{2i})}} \\ \boldsymbol{LWL = \mu_0 - L_1\sigma\sqrt{\frac{\lambda}{n_i(2-\lambda)}(1-(1-\lambda)^{2i})}} \end{cases} \quad \textbf{(II.26)}$$

where, L_1 is the warning coefficient of VSS-EWMA.

The area between LWL and UWL is called as safety zone. The area between LCL and LWL, as well as the area between UWL and UCL are called warning zones.

Let n_1 , n_2 be the smaller number of samples and larger number of samples, respectively. Hence, the smaller sample size (n_1) is taken in i^{th} sample if Z_{i-1} falls in safety zone and larger sample size (n_2) is taken if Z_{i-1} falls in warning zones.

$$\begin{cases} LWL < \mathbf{Z_{i-1}} < UWL \rightarrow \mathbf{n_i = n_1} \\ UWL < \mathbf{Z_{i-1}} < UCL \rightarrow \mathbf{n_i = n_2} \\ LWL < \mathbf{Z_{i-1}} < LCL \rightarrow \mathbf{n_i = n_2} \end{cases} \quad \textbf{(II.27)}$$

The determination of sample sizes in VSS-EWMA is critical for its effectiveness. Smaller and larger sample sizes are chosen based on the process's current position relative to the control limits **[49]**. Specifically, smaller sample sizes are used when the process is within the safety zone, indicating stability, while larger sample sizes are employed when the process enters the warning zones, signifying potential shifts or disturbances.

The EWMA statistic with variable sample size follows the same formulation as the traditional EWMA statistic and $\boldsymbol{Z_0 = \mu_0}$, incorporating the current process measurement and the previous EWMA statistic **[49]**. This statistic serves as a dynamic measure of process performance, reflecting recent process behavior while considering historical data.

The calculation of control limits for VSS-EWMA involves modifications to accommodate the variability in sample sizes **[49]**. These control limits are derived using

equations that integrate the process mean, standard deviation, and the smoothing parameter, ensuring that they adapt to changes in sample sizes and process conditions.

The control limits of VSS-EWMA are modified as below:

$$\begin{cases} UCL_i = \mu_0 + L_2\sigma\sqrt{\frac{\lambda}{n_i(2-\lambda)}(1-(1-\lambda)^{2i})} \\ CL = \mu_0 \\ LCL_i = \mu_0 - L_2\sigma\sqrt{\frac{\lambda}{n_i(2-\lambda)}(1-(1-\lambda)^{2i})} \end{cases} \quad \textbf{(II.28)}$$

For i=0,$Z_0 = \mu_0$ and UCL_0, LCL_0 are starting value of control limits and can be determined as follows:

$$\begin{cases} UCL_0 = \mu_0 + L_2\sigma\sqrt{\frac{\lambda}{n_1(2-\lambda)}} \\ CL = \mu_0 \\ LCL_0 = \mu_0 - L_2\sigma\sqrt{\frac{\lambda}{n_1(2-\lambda)}} \end{cases} \quad \textbf{(II.29)}$$

Since $Z_0 = \mu_0$ is used as a starting point in the traditional VSS control charts and the center line (μ_0) is in the safe zone, smaller sample size is used as first number of samples in this method. This is why n_1 is used as starting sample size in Equation (**II.29**). In addition, in some papers in the literature of VSS control charts such as Flaig (1991), the smaller sample size is chosen as the initial sample size. The sample sizes for next samples **(n_i i = 2, 3, ...)** is calculated by Equation (**II.27**).

In the initial stage, the starting values for control limits are determined based on the process mean and standard deviation, similar to FSS-EWMA. However, in VSS-EWMA, adjustments are made to these starting values to reflect the initial sample size and ensure compatibility with the variable sample size methodology **[39]**.

The effectiveness of VSS-EWMA lies in its ability to provide timely detection of process shifts and disturbances while maintaining stability in performance evaluation. By dynamically adjusting sample sizes and control limits, VSS-EWMA charts offer improved sensitivity to small shifts in process parameters, making them valuable tools in industries where continuous monitoring of process performance is crucial for maintaining product quality and efficiency **[39]**.

II.14. VSSILF-EWMA (Variable Sample Size EWMA as an Integer Linear Function)

The VSSILF-EWMA control chart is a sophisticated method proposed by Amiri, Nedaie, and Alikhani to enhance statistical process control. It introduces an adaptive approach to monitoring process parameters, particularly focusing on improving the detection of small shifts **[49]**.

The proposed VSSILF-EWMA model introduces a new approach for determining sample sizes in VSS-EWMA control charts. In this model, the sample size is determined as an integer linear function of the EWMA statistic value. The relationship between sample size and EWMA statistic follows a linear function, ensuring adaptability to changes in process conditions **[50].**

Assume that a process follows normal distributions, in which we can say the process is in-control (IC) if $\boldsymbol{X_i \sim N(0, \sigma^2)}$ and the process is out-of-control (OC) if $\boldsymbol{X_i \sim N(\mu_0, \sigma^2)}$.
In this model we focus on the mean shifts. Hence, we suppose $\boldsymbol{\mu_0 \neq 0}$.

In determining the sample size for the VSS-EWMA control chart, we utilize an integer linear function based on the EWMA statistic value (VSSILF-EWMA). This entails establishing a linear relationship between the EWMA statistic and the sample size. Notably, we allocate smaller sample sizes for subsequent samples when the current $\boldsymbol{Z}$ value is closer to the center line, and larger sizes otherwise. Specifically, the minimum sample size ($\boldsymbol{n_1}$) is assigned when $\boldsymbol{Z}$ equals the center line ($\boldsymbol{CL}$), while the maximum sample size ($\boldsymbol{n_2}$) is allotted when $\boldsymbol{Z}$ equals either the upper control limit ($\boldsymbol{UCL}$) or the lower control limit ($\boldsymbol{LCL}$). Figure (**II.4**) illustrates the relationship between sample size ($\boldsymbol{n}$) and statistic value ($\boldsymbol{Z}$). Consequently, we define the sample size function as follows **[49]**:

$$\boldsymbol{n_i = f(|Z_{i-1}|)} \qquad \textbf{(II.30)}$$

The relation between sample size and Z is linear. Also, the value of sample size must be integer. So, an integer linear function can be defined for sample size.

$$\boldsymbol{n_i = \left[\left(\frac{n_2 - n_1}{UCL_{i-1} - \mu}\right)|Z_{i-1}| + \left(n_1 - \frac{n_2 - n_1}{UCL_{i-1} - \mu} \times \mu\right)\right]} \qquad \textbf{(II.31)}$$

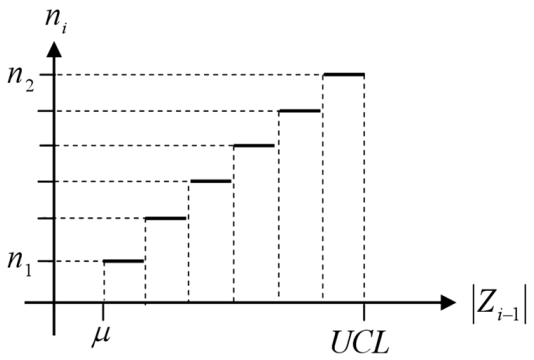


Figure II.4: Sample size function plot as statistic value

The sample sizes (n_i) specified in Equation (**II.31**) represent positive values ranging between n_1 and n_2. As previously mentioned, n_1 and n_2 denote the minimum and maximum sample sizes we aim to employ, respectively. UCL_{i-1} denotes the upper control limit of the proposed chart. μ represents the mean of the process under control, while $|Z_{i-1}|$ denotes the statistical value for the $(i-1)^{th}$ sample.

Given the variable nature of the sample size, the control limits also vary accordingly. Upon substituting Equation **(II.25)**, the upper and lower control limits of this chart can be expressed as the equation (**II.31**):

$$\begin{cases} UCL_i = \mu_0 + L_3\sigma\sqrt{\frac{\lambda}{|n_i|(2-\lambda)}(1-(1-\lambda)^{2i})} \\ CL = \mu_0 \\ LCL_i = \mu_0 - L_3\sigma\sqrt{\frac{\lambda}{|n_i|(2-\lambda)}(1-(1-\lambda)^{2i})} \end{cases} \quad \textbf{(II.32)}$$

For $i=0$, $Z_0 = \mu_0$, $n_0 = n_1$, and UCL_0, LCL_0 are starting value of control limits and can be determined as follows:

$$\begin{cases} UCL_0 = \mu_0 + L_3\sigma\sqrt{\frac{\lambda}{n_1(2-\lambda)}} \\ CL = \mu_0 \\ LCL_0 = \mu_0 - L_3\sigma\sqrt{\frac{\lambda}{n_1(2-\lambda)}} \end{cases} \quad \textbf{(II.33)}$$

The calculation of sample size (n), upper and lower control limits (UCL and LCL), and statistical value (Z) is necessary for each sample. The VSS-EWMA, functioning as an integer linear function (VSSILF-EWMA), triggers an alert when Z exceeds UCL or falls below LCL. It's important to note that our proposed method extends the traditional VSS approach. Hence, we initialize Z as the baseline for the EWMA statistic and n_1 (the smaller sample size) as the initial sample size **[49]**.

II.15. Adaptive Sliding Window EWMA (ASW-EWMA) Control Chart

The ASW-EWMA control chart introduces a dynamic approach to process monitoring and anomaly detection.

The fundamental principle remains rooted in EWMA methodology, where Exponentially Weighted Moving Average (EWMA) control charts are employed to track process stability over time by assigning exponentially decreasing weights to historical data points while giving more weight to recent observations.

The operation of the ASW-EWMA control chart involves the continual recalculation of control limits (***LCL***/***UCL***) based on the EWMA statistic, incorporating both historical and incoming data. Samples flagged as out of control values are considered potential anomalies, prompting further investigation or action. Conversely, normal samples are added to the training set, contributing to the ongoing refinement of control limits. The calculating of the control limits as the same of static EWMA.

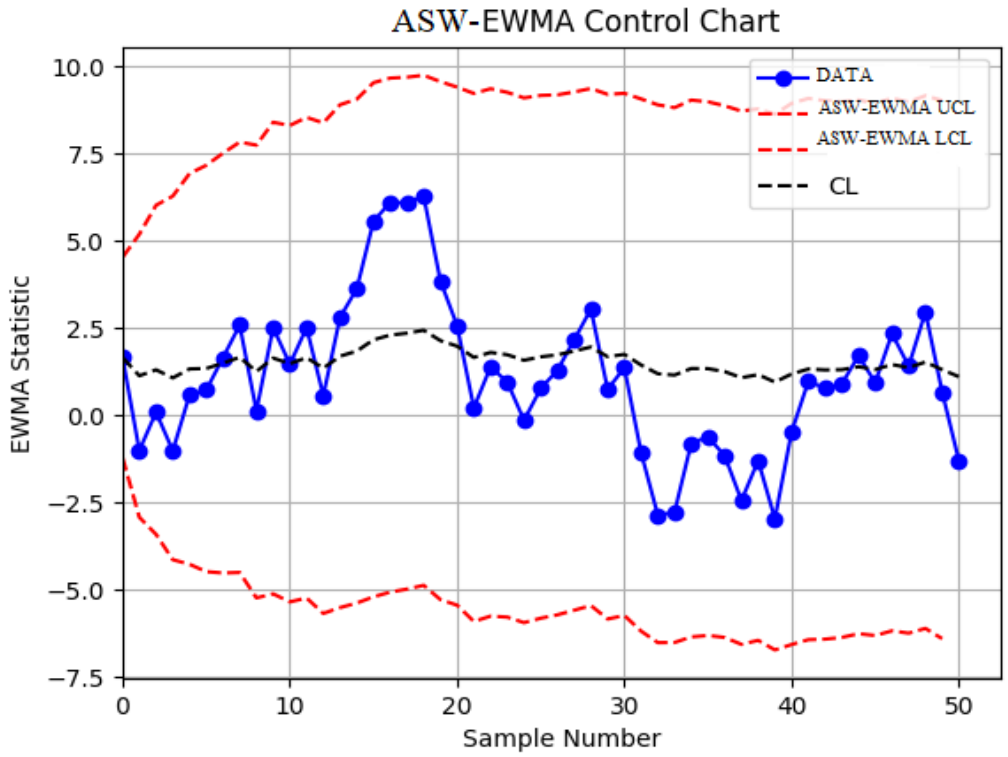


Figure II.5: ASW-EWMA Control Chart

This adaptive approach ensures that the control chart evolves with the changing process dynamics, effectively adapting to emerging patterns and anomalies. By dynamically adjusting

to new information, the model enhances its responsiveness to both gradual shifts and sudden disturbances in the process.

II.16. Conclusion

In conclusion, while traditional control charts like Shewhart, CUSUM, and EWMA are effective in detecting process changes, the advent of Adaptive EWMA charts represents a significant advancement. In cyber-security, Adaptive EWMA charts dynamically adjust to evolving threats, offering real-time detection of cyber-attacks. Their adaptability ensures timely identification and mitigation of subtle network traffic variations, thereby enhancing defenses against malicious incursions.

Adaptive EWMA charts operate by adjusting control limits and smoothing parameters in response to process changes, making them highly effective against dynamic cyber threats. Different models, such as VSS-EWMA, VSSILF-EWMA, and ASW-EWMA, have been developed to improve detection capabilities. Embracing these adaptive models is crucial for staying ahead of adversaries and protecting critical digital assets.

In Chapter Three, we will study the comparison of control limits when parameters are changed for each chart and simulate the performance of these control charts for detecting DoS and DDoS attacks. This comparative analysis will highlight the strengths and weaknesses of each model in various scenarios, providing valuable insights into their practical application in cyber-security.

Chapter III

Cyber-Attacks Detection Using Adaptive EWMA based models

Chapter III Cyber-Attacks Detection Using Adaptive EWMA based models

III.1. Introduction

In this chapter, we delve into the efficacy of some Adaptive EWMA control chart models in detecting DoS and DDoS attacks within IP networks through Python-based simulations. Our objective is to assess their effectiveness across diverse attack scenarios, particularly focusing on TCP SYN flood and Smurf attacks. By conducting a comprehensive evaluation under varying conditions, we aim to provide insights into the comparative performance of these charts. Leveraging network traffic data sourced from the DARPA99 dataset, we embark on a comparative analysis to gauge their respective capabilities in safeguarding network security.

III.2. Detecting DoS and DDoS Cyber-Attacks Using AEWMA based models

Detecting DoS and DDoS attacks within a network involves monitoring various network traffic parameters, such as messages, segments, bits, protocols, ports, and IP addresses. These parameters often experience significant fluctuations during such malicious activities.

Adaptive EWMA control charts, including FSS-EWMA, VSS-EWMA, VSSILF-EWMA, and ASW-EWMA, are effective tools for identifying deviations from the norm and detecting abnormal traffic patterns associated with DoS and DDoS attacks.

Among these, the Adaptive EWMA approaches offer a dynamic and responsive method for monitoring network traffic. FSS-EWMA rapidly adapts to changes, while VSS-EWMA adjusts sampling frequency based on traffic variability. VSSILF-EWMA introduces interleaving to enhance sensitivity, and ASW-EWMA dynamically weights samples for improved detection accuracy. By leveraging these adaptive techniques, network administrators can effectively identify subtle and substantial shifts in monitored parameters indicative of DoS and DDoS attacks.

The detection process involves several key steps:

1. Capturing and Collecting Network Traffic: Utilizing tools like TCPDUMP, Wireshark, SolarWinds Deep Packet Inspection and Analysis tool, etc., enables the capture of network

traffic reflecting network activities. These tools can be strategically placed across the network, facilitating real-time traffic recovery for subsequent analysis, including the detection of anomalies, intrusions, and potential DoS and DDoS attacks.

2. Preprocessing Data and Extracting Parameters: Given that DoS and DDoS attacks can impact various network traffic parameters, it's crucial to identify and isolate the parameters for monitoring. Since network traffic is typically captured in raw form, preprocessing becomes necessary. This step involves extracting and isolating detection parameters and transforming them into measurable input data suitable for control charts.

3. Constructing Control Charts: Learning from normal traffic data aids in calculating control limits (CL/UCL/LCL) for EWMA, VSS-EWMA, VSSILF-EWMA, and the ASW-EWMA charts. These limits delineate the acceptable range of parameter variation under normal network operation, absent targeted attacks.

4. Detecting and Identifying Attacks: Test data, potentially containing DoS and DDoS attacks, is used to compute characteristic statistics for each chart. Comparing sample values with control limits determines whether monitored traffic is normal or abnormal. If characteristics fall within the UCL and LCL bounds, the traffic is deemed normal. However, if they exceed these limits, an abnormal traffic pattern triggers a DoS or DDoS attack detection alarm, providing detailed information regarding the attack type, victim, and occurrence date.

This integrated approach, outlined in Figure III.1, showcases the general procedure for utilizing AEWMA based models control charts for detecting DoS and DDoS attacks, ensuring robust network security.

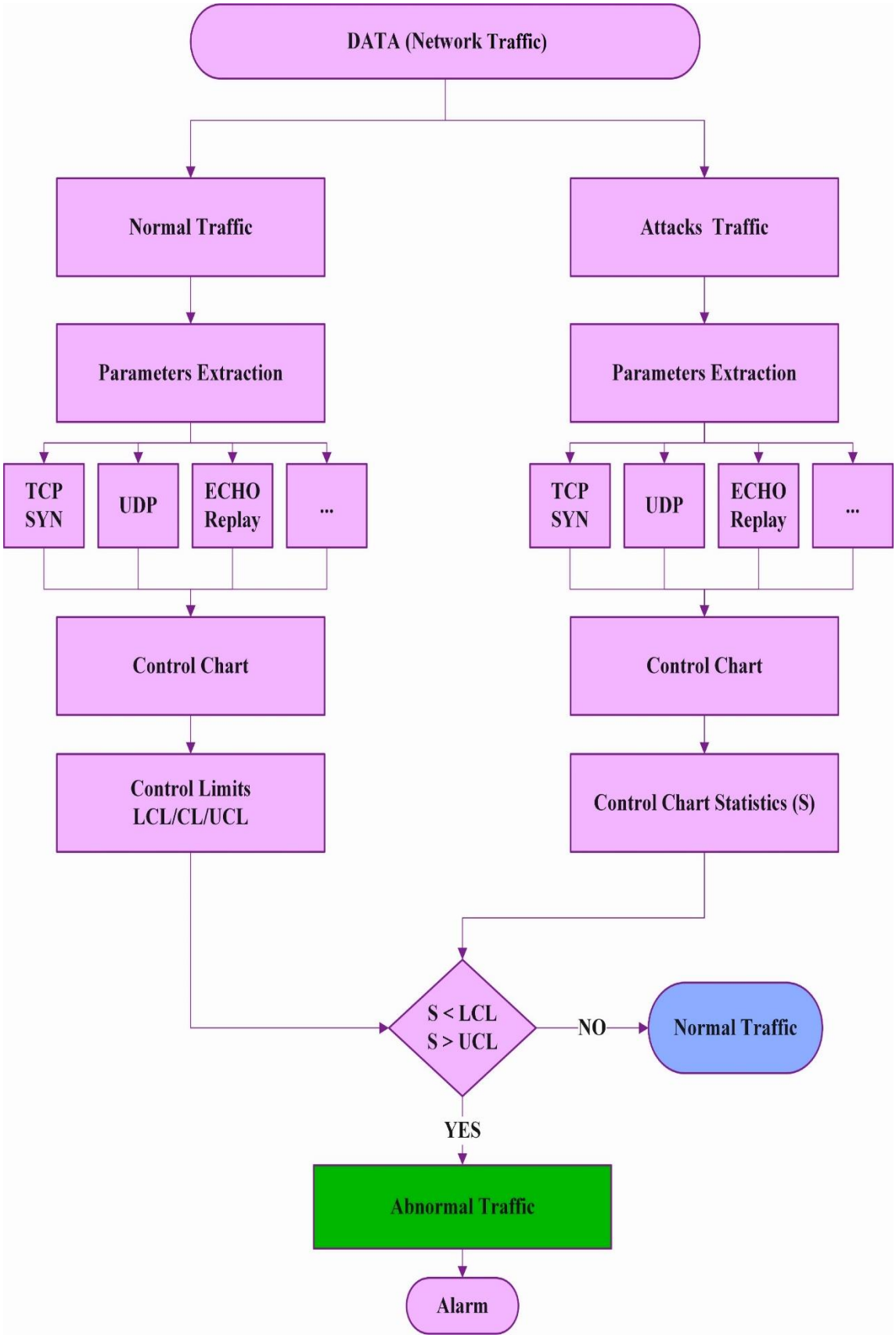


Figure III.1: Process of DoS and DDoS Attacks Detection Using Control Charts

III.3. Overview of the DARPA99 Dataset

The DARPA99 dataset, also known as IDEVAL (Intrusion Detection Evaluation), stands as a cornerstone in evaluating intrusion detection systems, encompassing a vast array of network traffic data meticulously curated by the Lincoln Laboratory of the Massachusetts Institute of Technology (MIT), with support from DARPA and the Air Force Research Laboratory (AFRL). Representing traffic captured from a simulated real-world network akin to that of a military base connected to the internet, this dataset serves as a pivotal resource for assessing intrusion detection mechanisms.

Comprising a simulated environment mirroring a military base network, the dataset's topology delineates two distinct realms: the internal network, simulated by machines on the left, and the external internet, emulated by machines on the right. These include various servers, such as Pascal, Zeno, Marx, Hume, Kant, and Aesop, each equipped with specific functionalities and operating systems, alongside virtual hosts facilitating IP address impersonation.

Incorporating diverse attack scenarios, including Denial-of-Service (DoS) attacks, the DARPA99 dataset captures the intricate interplay of network activities. Two sniffer machines, Locke and Solomon, running Solaris 2.6, were strategically positioned to capture internal and external network traffic, respectively, utilizing UNIX TCPDUMP for data collection. With a comprehensive collection spanning five weeks, comprising both normal and attack traffic, the dataset offers invaluable insights into intrusion detection system evaluation.

Moreover, the dataset's meticulous curation extends to its partitioning, with three weeks allocated for training data and two weeks designated for testing purposes. The training set provides a foundation devoid of attacks in its initial and final weeks, with the middle week incorporating simulated attacks. Conversely, the test set presents a diverse array of attack categories for rigorous system evaluation.

As a widely recognized benchmark in the field of intrusion detection, the DARPA99 dataset facilitates standardized evaluation methodologies, enabling researchers to benchmark and refine intrusion detection systems effectively. Its availability has catalyzed advancements in network security, contributing significantly to the evolution of intrusion detection technologies over time.

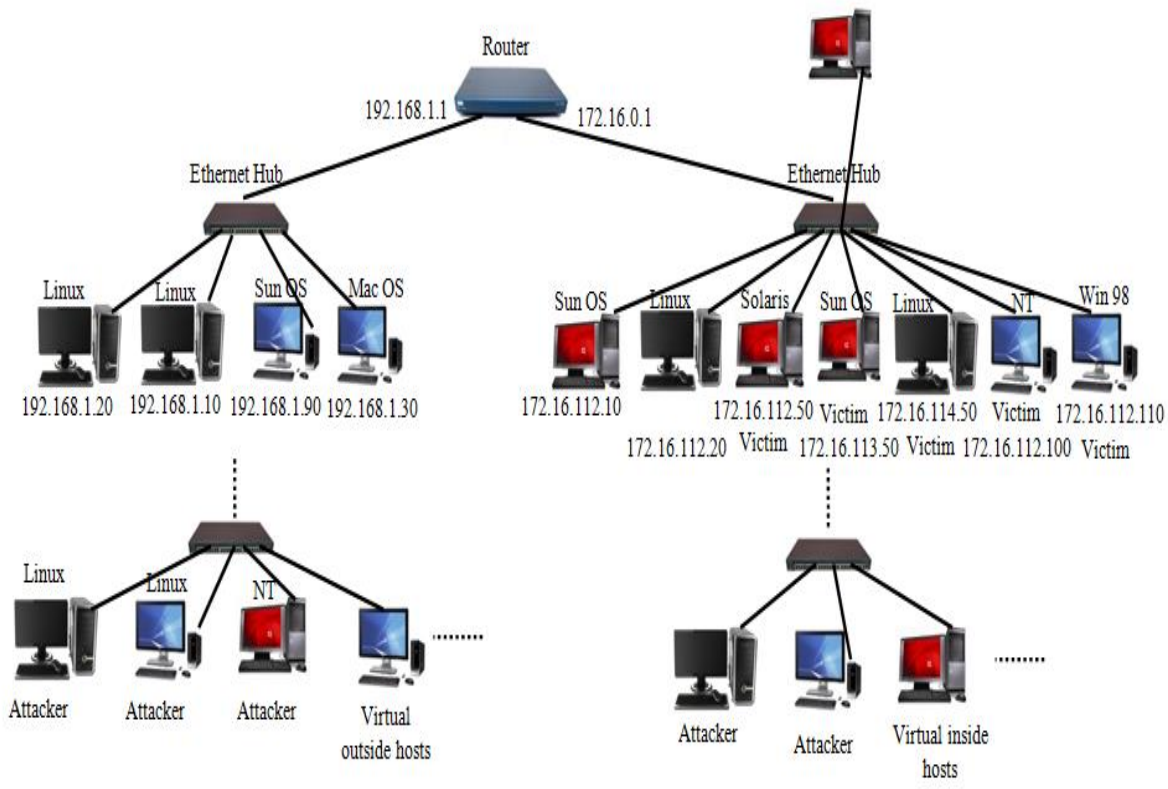


Figure III.2: Network Topology used in DARPA Simulation

III.4. Data Preprocessing

Given the comprehensive nature of the DARPA99 dataset, encompassing diverse network traffic parameters such as frames, packets, protocols, and messages (refer to Figure III.3), preprocessing becomes imperative to extract pertinent detection parameters and tailor them for integration into control charts. Notably, as TCP SYN flood, UDP flood, and Smurf DoS and DDoS attacks predominantly hinge on the voluminous transmission of SYN segments, UDP datagrams, and ICMP Echo reply messages respectively, preprocessing entails the meticulous extraction, manipulation, and organization of these distinct data structures.

Utilizing Wireshark, a network analyzer, on the DARPA99 TCP UDMP files facilitated the isolation of relevant parameters, specifically SYN segments, UDP datagrams, and ICMP Echo reply messages. This filtration process culminated in the creation of files exclusively housing these critical parameters, as depicted in Figure III.4, exemplifying the filtration of TCP SYN segments. Subsequently, leveraging a combination of MySQL and Java, we executed multifaceted operations on the filtered Wireshark files. These operations encompassed diverse tasks, including quantifying message counts per unit time or

measurement interval, categorizing them based on destination, among others. Moreover, this processing stage ensured the transformation of data into formats amenable for direct manipulation using computational tools such as Python, MATLAB, or analogous software. Figure III.5 illustrates the refined form of TCP SYN segments and ICMP Echo reply messages post-preprocessing, demonstrating the efficacy of this preparatory phase in streamlining data analysis and facilitating subsequent computational endeavors.

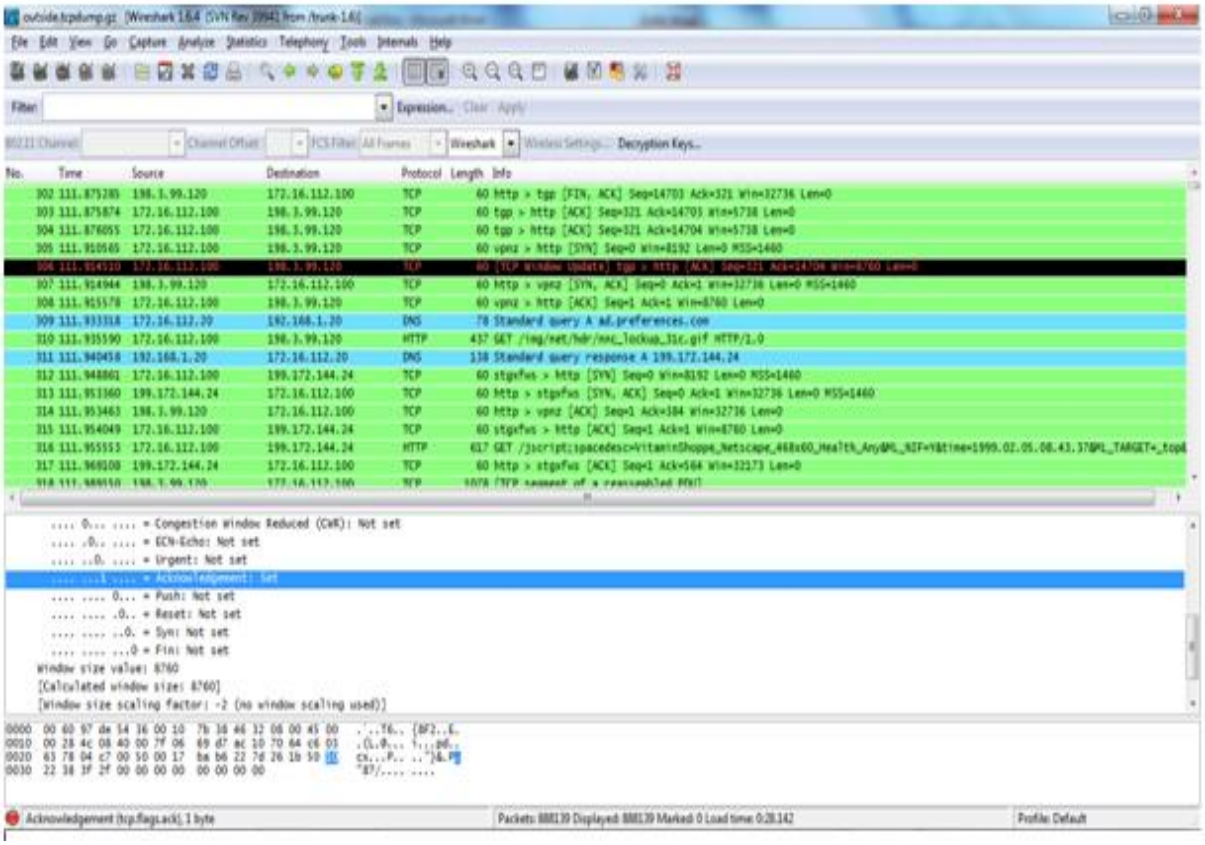


Figure III.3: Raw DARPA99 Traffic Visualized with Wireshark (example: week 2/day 3)

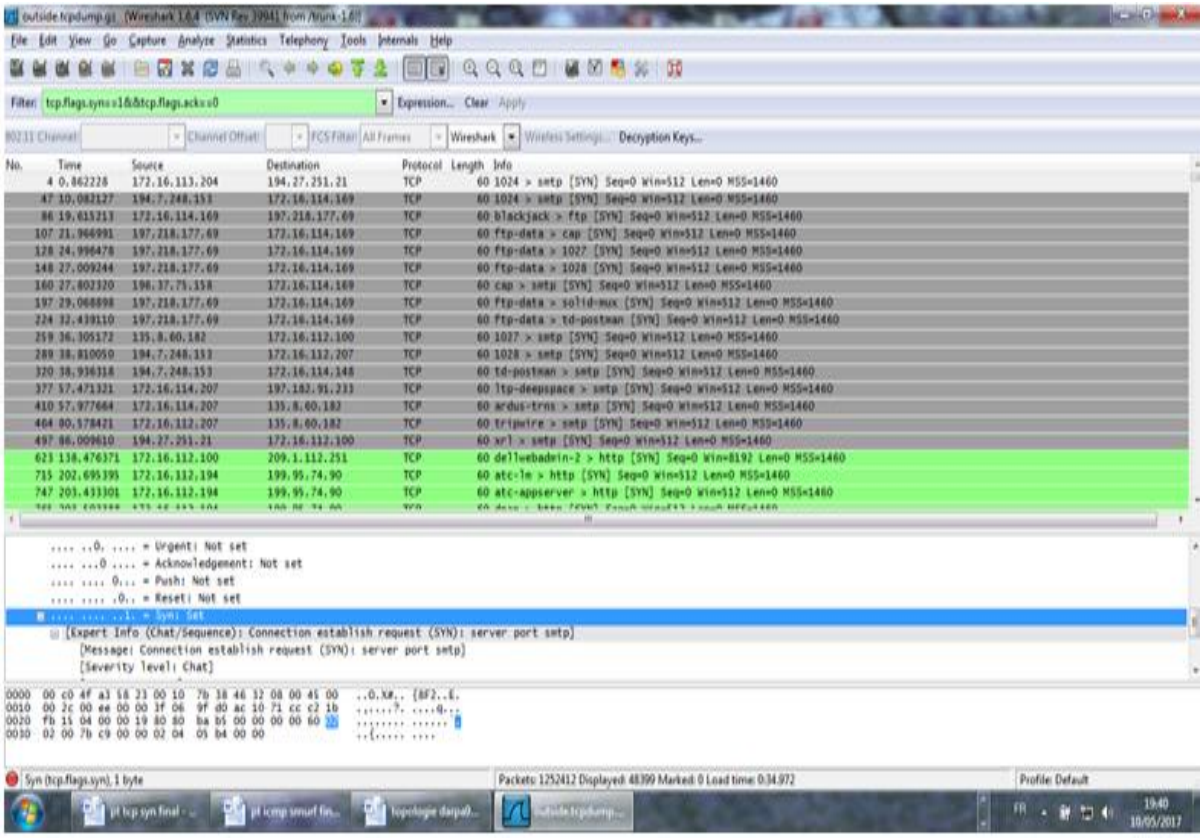
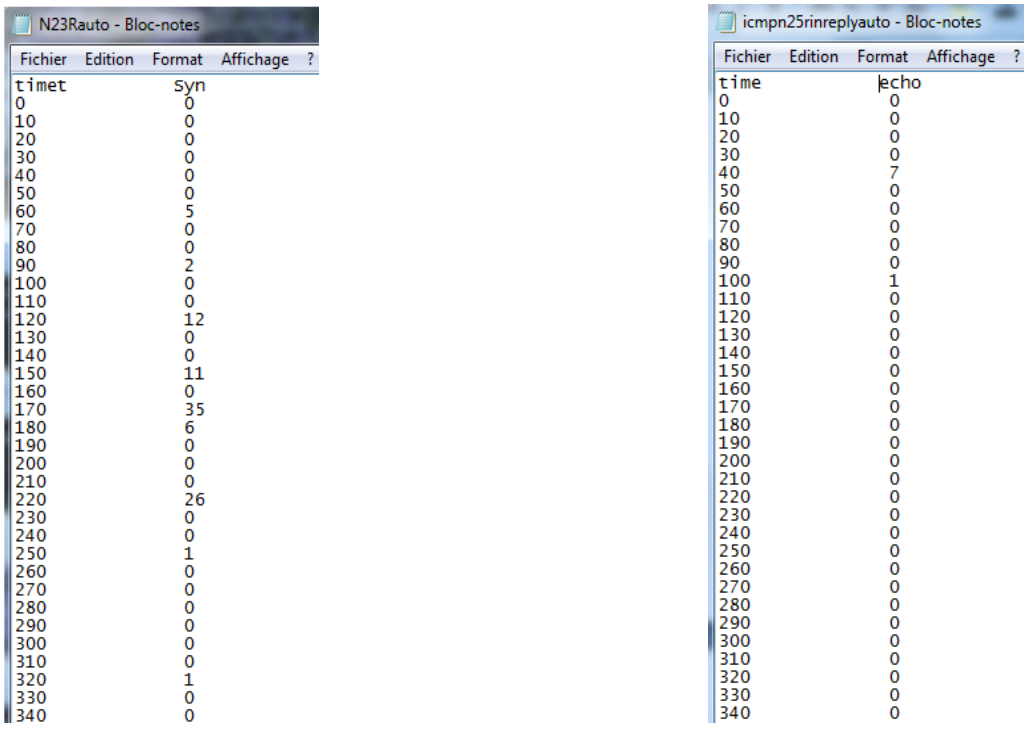


Figure III.4: Filtering of SYN Segments with Wireshark



N23Rauto - Bloc-notes

Fichier Edition Format Affichage ?

timet	Syn
0	0
10	0
20	0
30	0
40	0
50	0
60	5
70	0
80	0
90	2
100	0
110	0
120	12
130	0
140	0
150	11
160	0
170	35
180	6
190	0
200	0
210	0
220	26
230	0
240	0
250	1
260	0
270	0
280	0
290	0
300	0
310	0
320	1
330	0
340	0

icmpn25rinreplyauto - Bloc-notes

Fichier Edition Format Affichage ?

time	echo
0	0
10	0
20	0
30	0
40	7
50	0
60	0
70	0
80	0
90	0
100	1
110	0
120	0
130	0
140	0
150	0
160	0
170	0
180	0
190	0
200	0
210	0
220	0
230	0
240	0
250	0
260	0
270	0
280	0
290	0
300	0
310	0
320	0
330	0
340	0

a) TCP SYN Segments b) ICMP Echo Reply Messages

Figure III.5: Examples of Detection Parameters After Preprocessing

III.5. Comparison of Control Limits

In this section, we will examine how the control limits vary when adjusting the parameters λ and L for each control chart. We will analyze the control limits for various types of control charts, including EWMA, VSS-EWMA, VSSILF-EWMA, and ASW-EWMA, in the context of detecting SYN and Smurf attacks. By comparing the control limits across these methodologies, we aim to understand their effectiveness and sensitivity in identifying these specific types of network attacks. This analysis will provide insights into the optimal parameter settings for each control chart to enhance detection accuracy.

For the EWMA control chart:

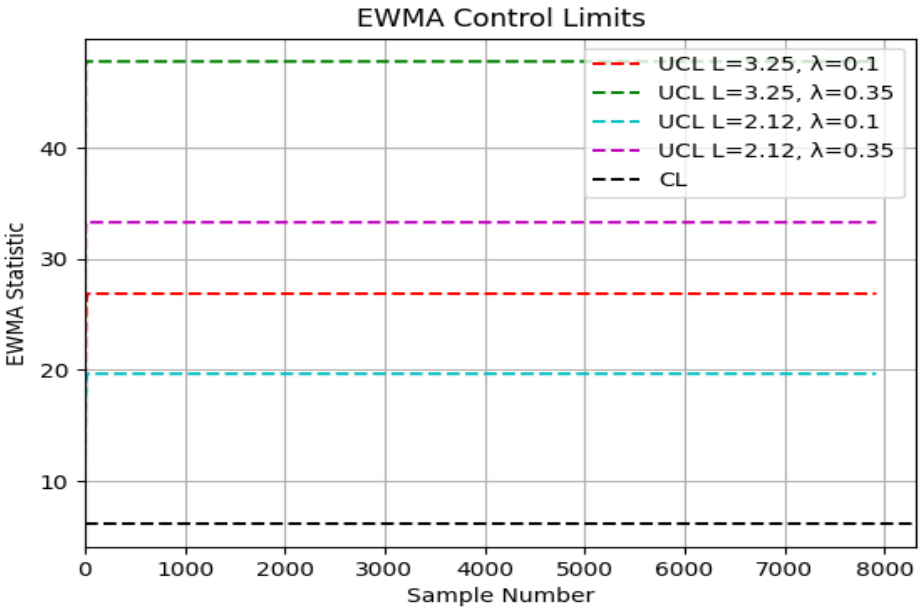


Figure III.6: Comparison of control limits for EWMA control chart using SYN attack

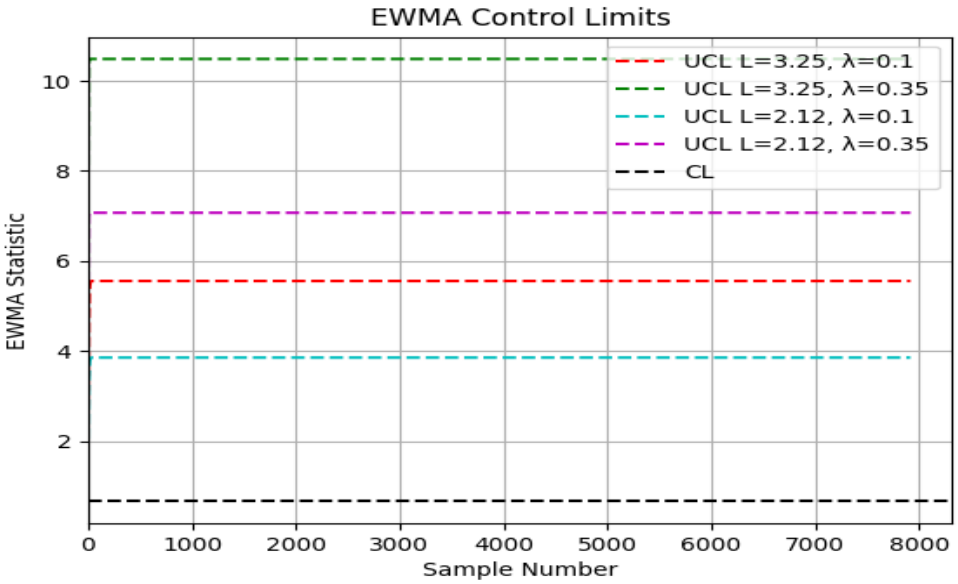


Figure III.7: Comparison of control limits for EWMA control chart using Smurf attack

For the VSS-EWMA control chart:

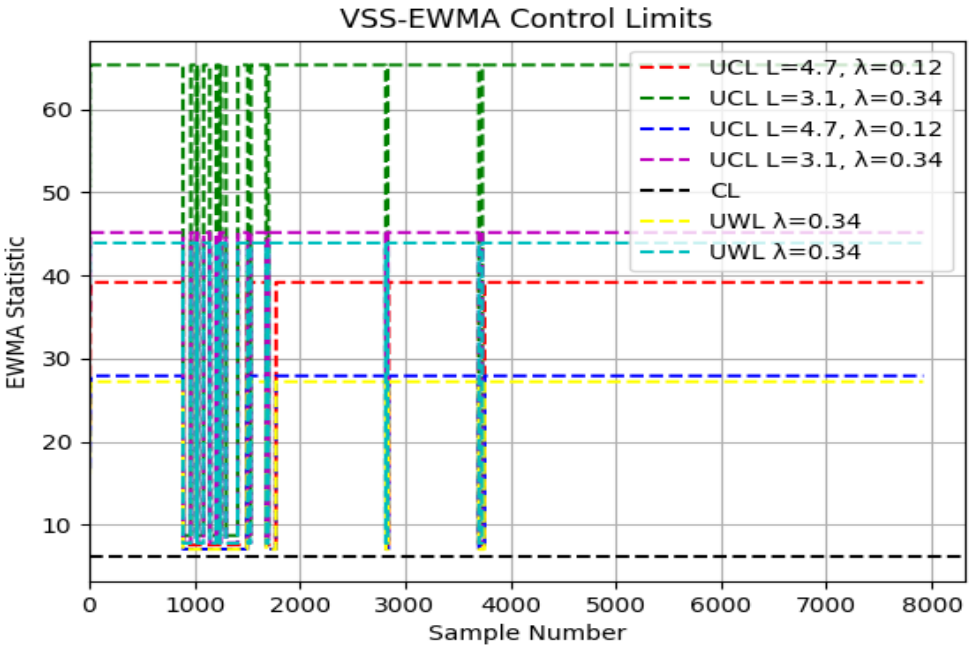


Figure III.8: Comparison of control limits for VSS-EWMA control chart using SYN attack

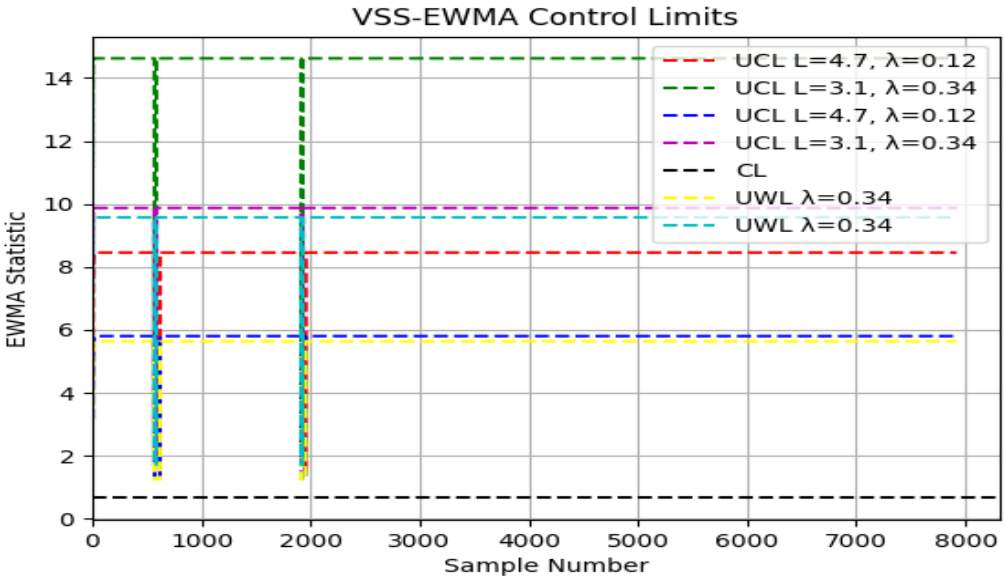


Figure III.9: Comparison of control limits for VSS-EWMA control chart using Smurf attack

For the VSSILF-EWMA control chart:

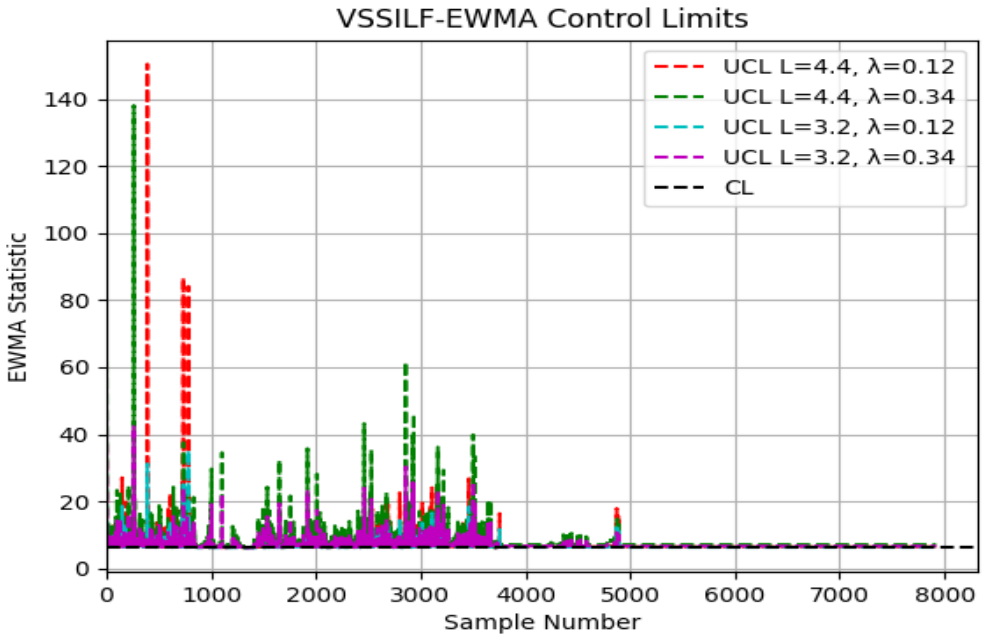


Figure III.10: Comparison of control limits for VSSILF-EWMA control chart using SYN attack

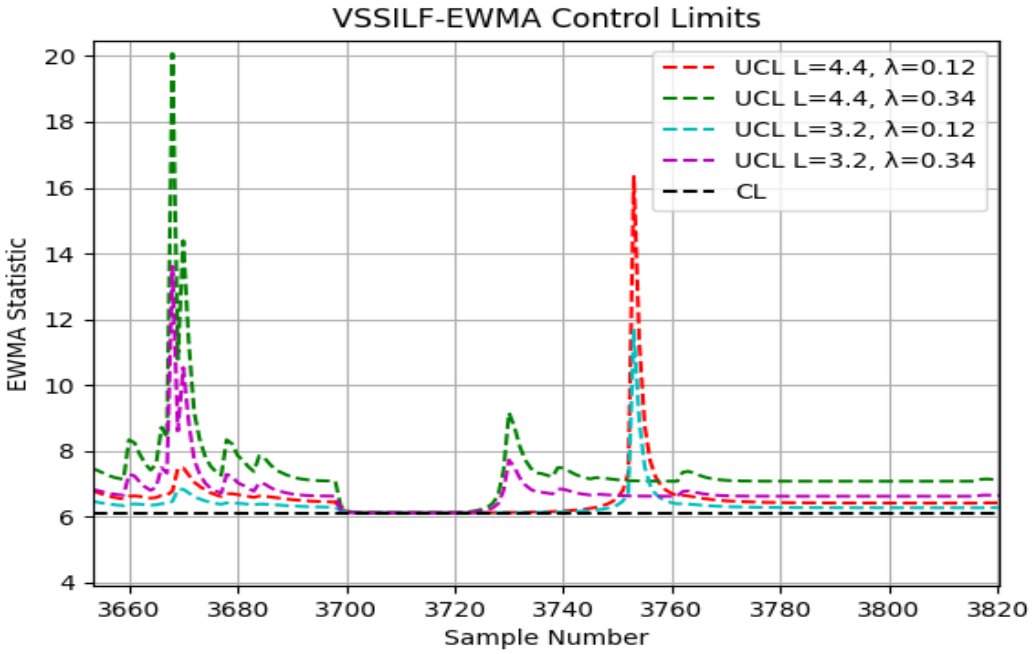


Figure III.11: Comparison of control limits for VSSILF-EWMA control chart using SYN attack with zoom

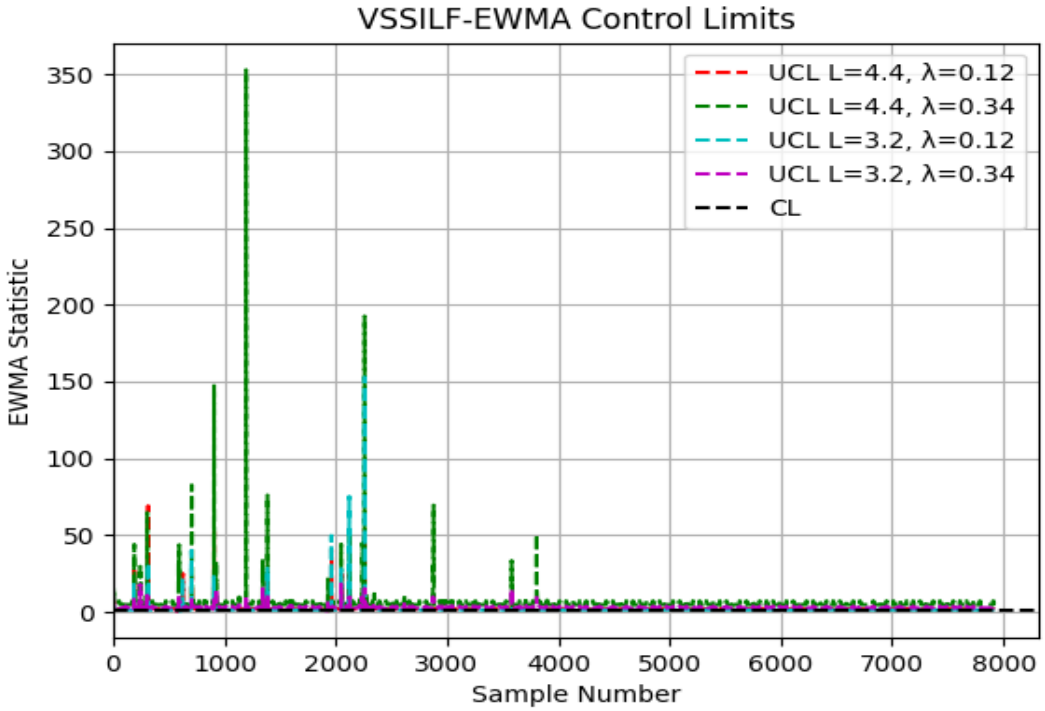


Figure III.12: Comparison of control limits for VSSILF-EWMA control chart using Smurf attack

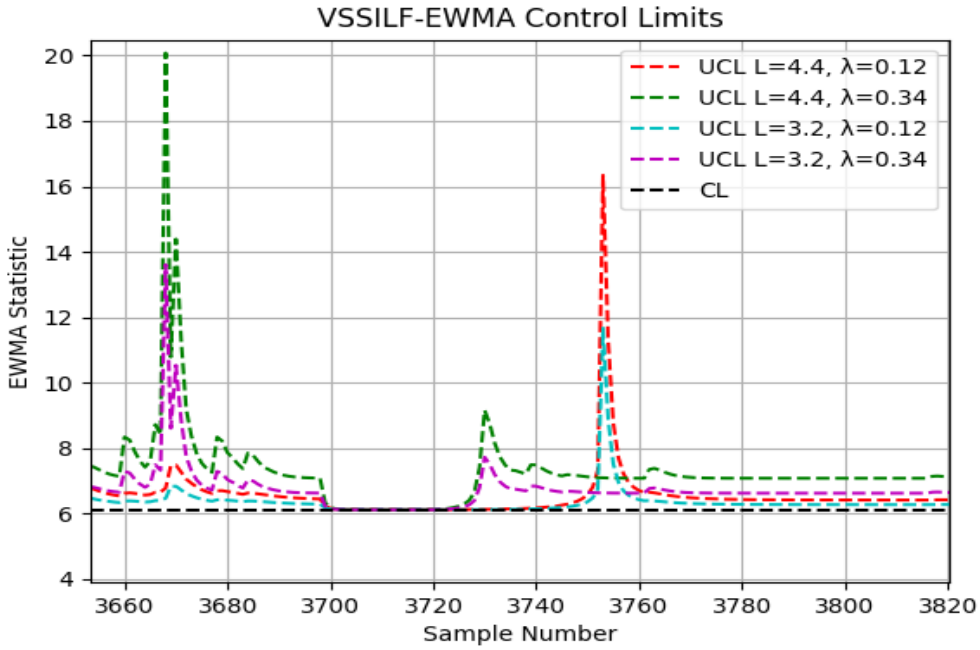


Figure III.13: Comparison of control limits for VSSILF-EWMA control chart using Smurf attack with zoom

For the ASW- EWMA control chart:

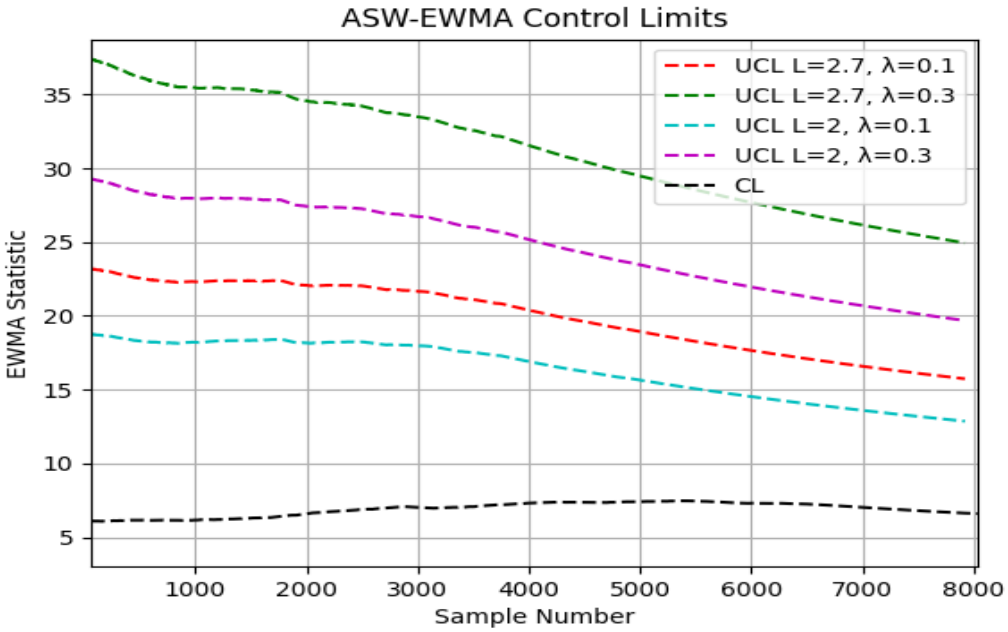


Figure III.14: Comparison of control limits for ASW-EWMA control chart using SYN attack

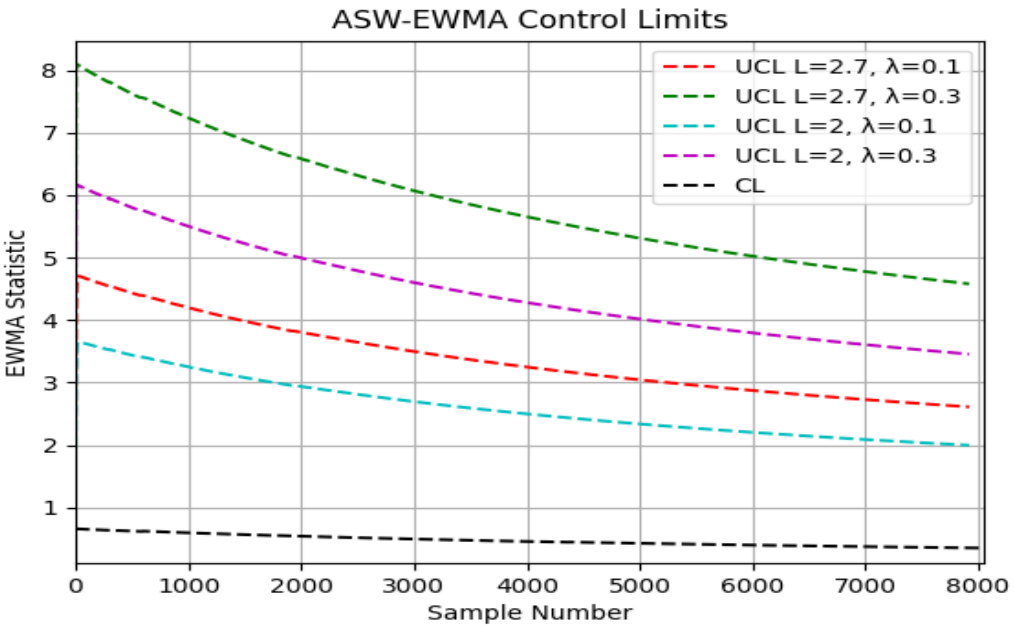


Figure III.15: Comparison of control limits for ASW-EWMA control chart using Smurf attack

III.6. Detection results using DARPA99 dataset

In this part, we'll look at spotting DoS and DDoS attacks, especially TCP SYN flood and Smurf attack, using data from DARPA99 dataset and Adaptive control charts-based models (like EWMA, VSS-EWMA, VSSILF-EWMA, and ASW-EWMA Control Chart). First, we'll pick out the key signs of each attack, such as SYN for TCP SYN flood and echo reply for Smurf. Then, we'll split the gathered traffic into two parts: regular traffic, which helps us set up control limits, and attack traffic, which we'll compare against those limits to see if things are going haywire. The attack traffic comes in three flavors: random traffic with lots of TCP SYN segments and echo replay messages, traffic with strong attacks, and traffic with weaker attacks. This way, we can see how well the Adaptive control charts handle different levels of TCP SYN flood and Smurf attacks.

III.6.1 TCP SYN flood Attack Detection

To examine the effectiveness of Adaptive EWMA control chart models (including VSS-EWMA, VSSILF-EWMA, and ASW-EWMA Control Chart) in detecting TCP SYN flood attacks, we'll evaluate their performance across three distinct scenarios: random traffic featuring TCP SYN segment flows, traffic subjected to high-intensity TCP SYN attacks, and

traffic experiencing low-intensity TCP SYN attacks. Our analysis will utilize training data sourced from the fifth day of the second week and testing data extracted from the second day of the fifth week of the dataset.

III.6.2. The Result of SYN Attack Traffic

III.6.2.1. Random Traffic of SYN Segments Flow

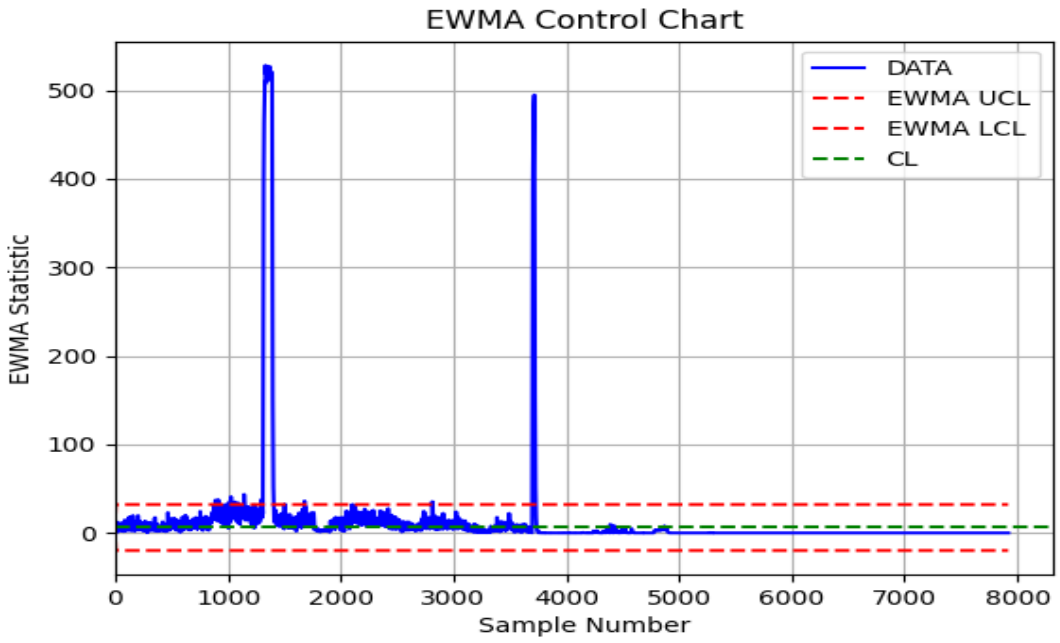


Figure III.16: EWMA Control Chart for the flow of SYN segments

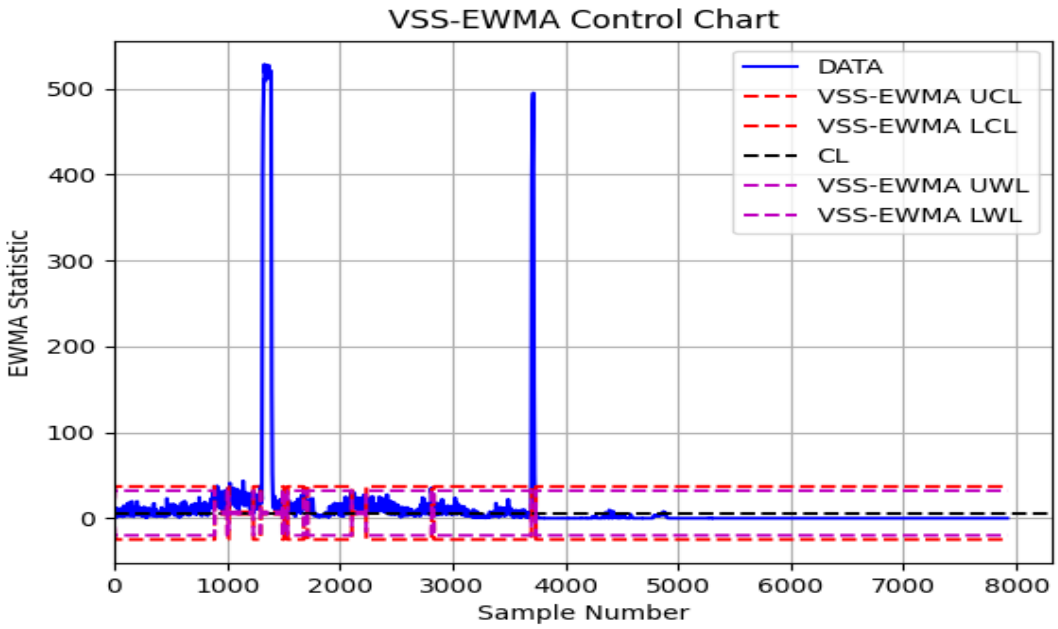


Figure III.17: VSS-EWMA Control Chart for the flow of SYN segments

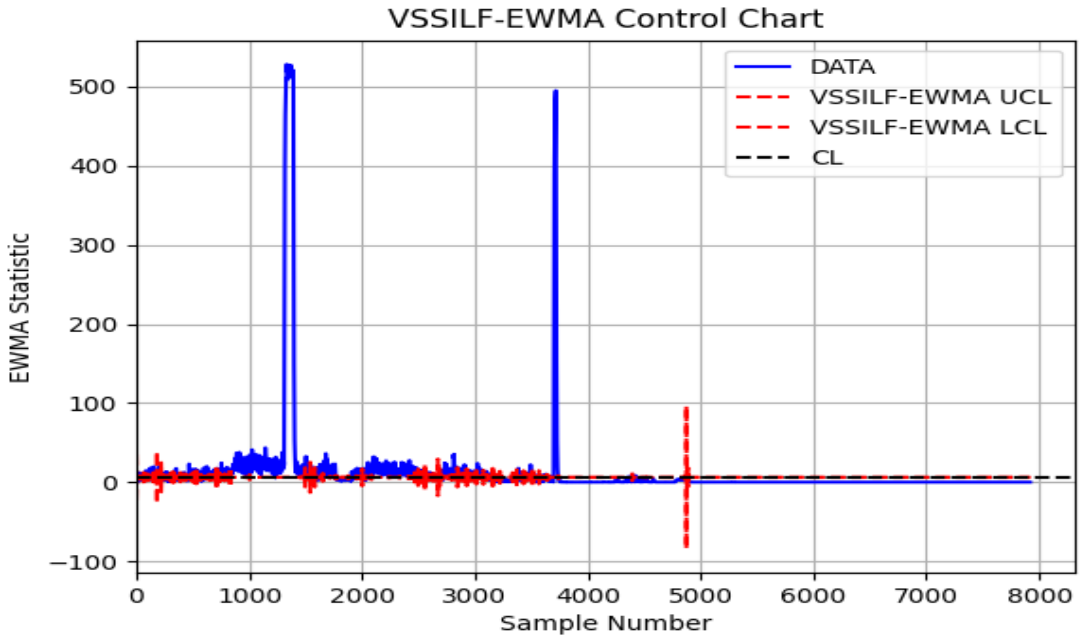


Figure III.18: VSSILF-EWMA Control Chart for the flow of SYN segments

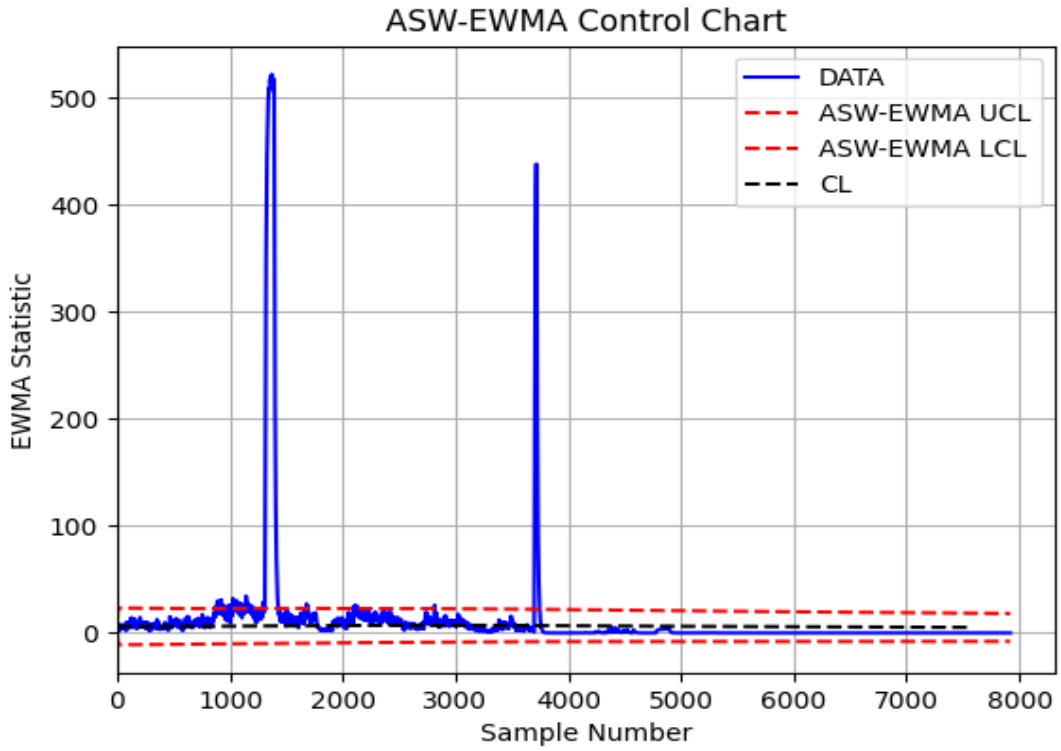


Figure III.19: ASW-EWMA Control Chart for the flow of SYN segments

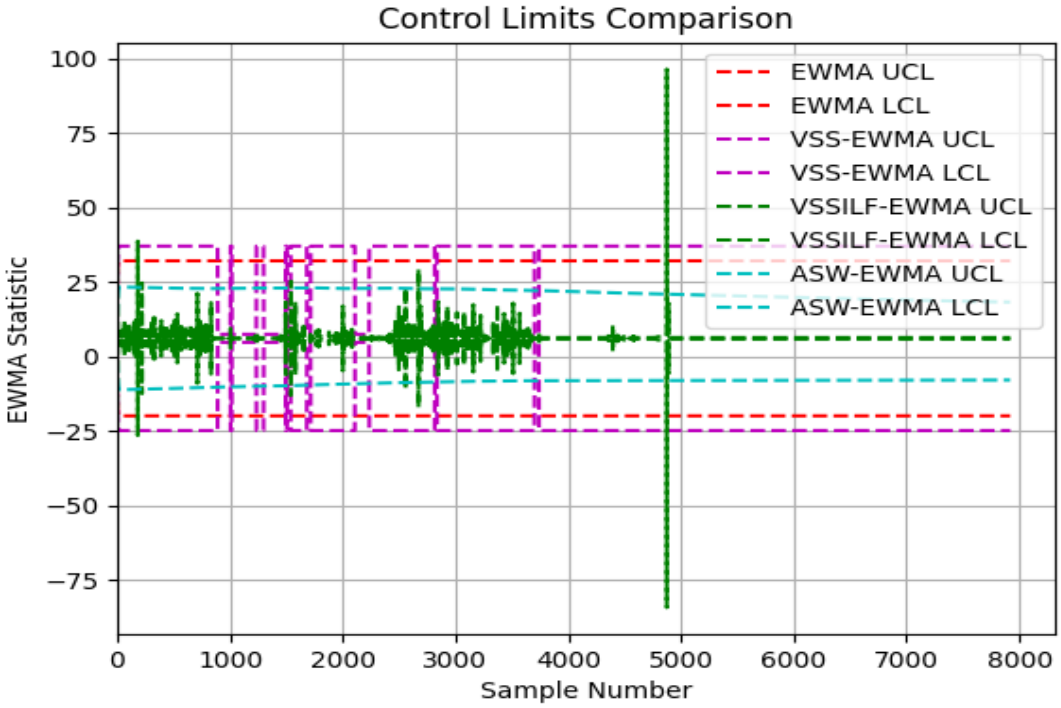


Figure III.20: Comparison of Control Charts Limits for SYN segments

III.6.2.2. Traffic with high intensity TCP SYN attacks

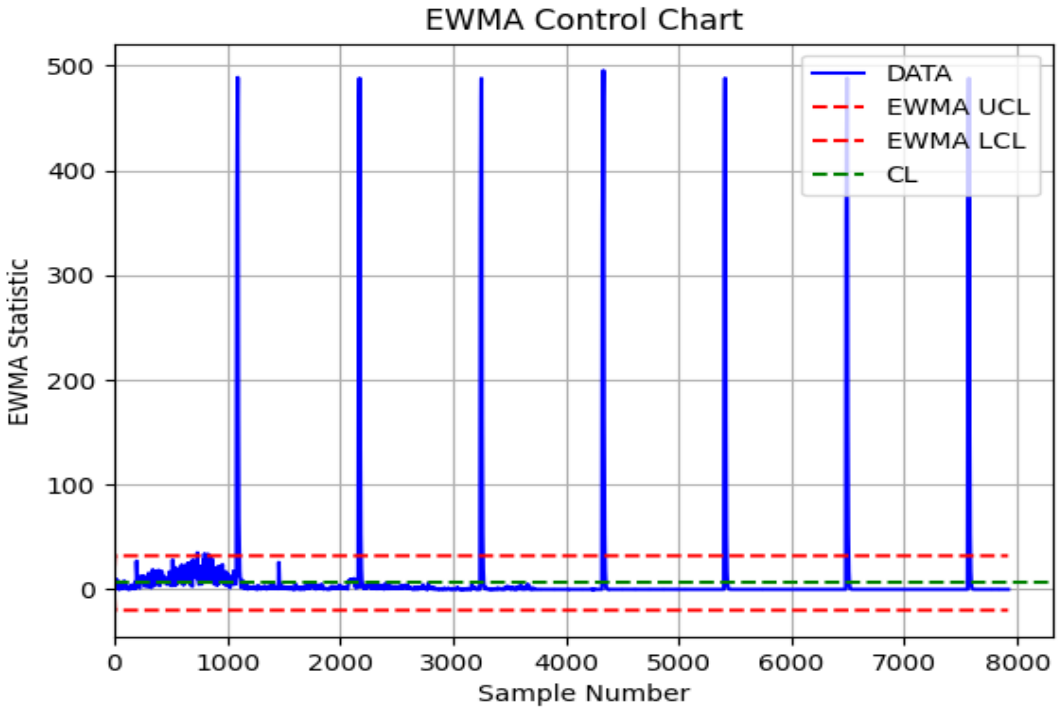


Figure III.21: EWMA Control Chart for high intensity of SYN segments

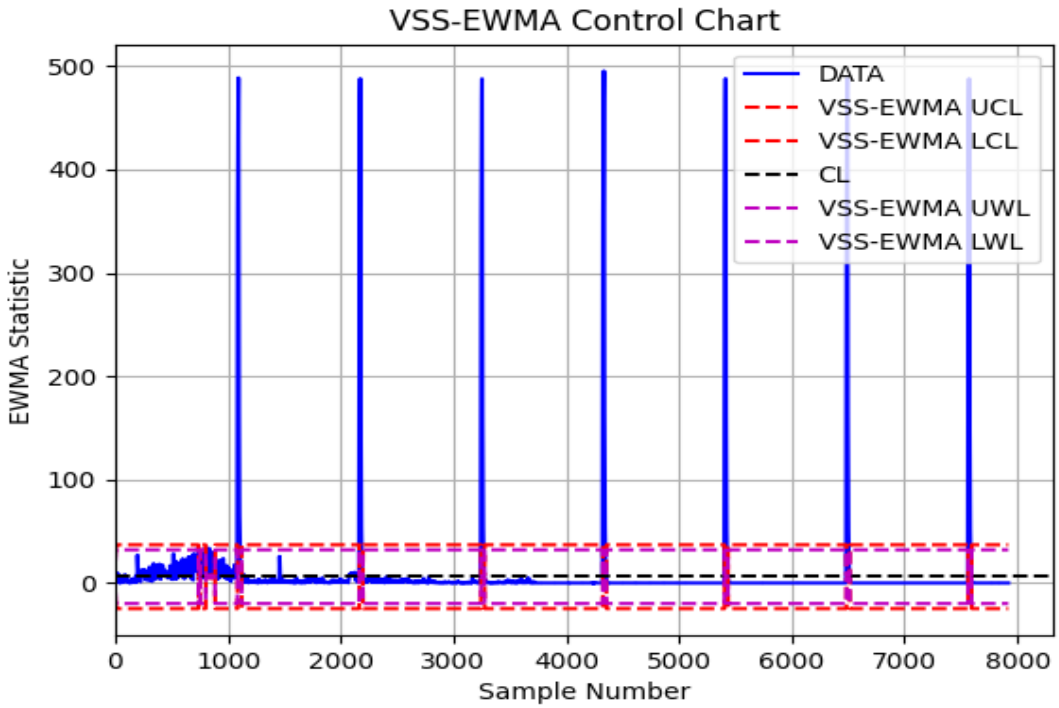


Figure III.22: VSS-EWMA Control Chart for high intensity of SYN segments

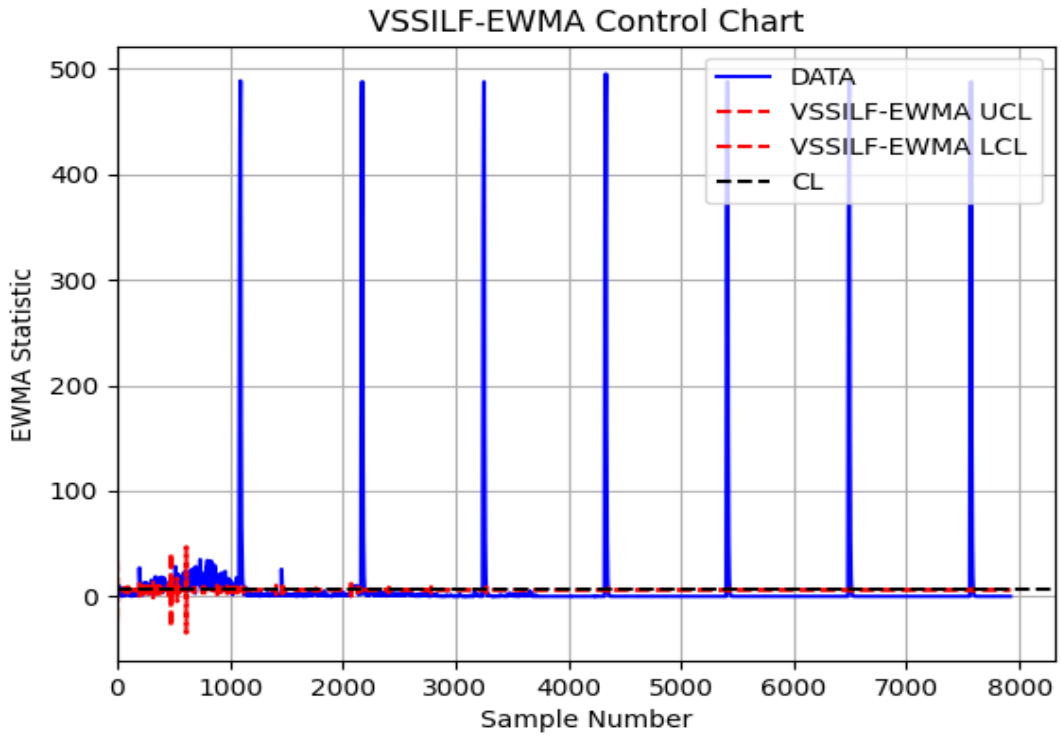


Figure III.23: VSSIF-EWMA Control Chart for high intensity of SYN segments

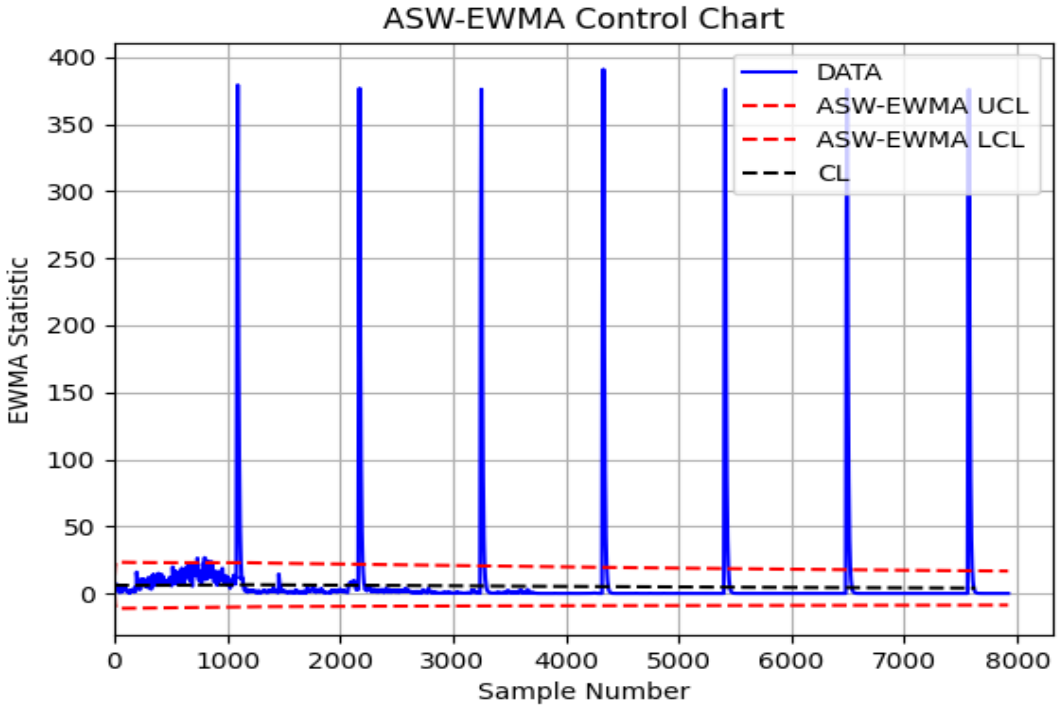


Figure III.24: ASW-EWMA Control Chart for high intensity of SYN segments

III.6.2.3. Traffic with low intensity TCP SYN attacks

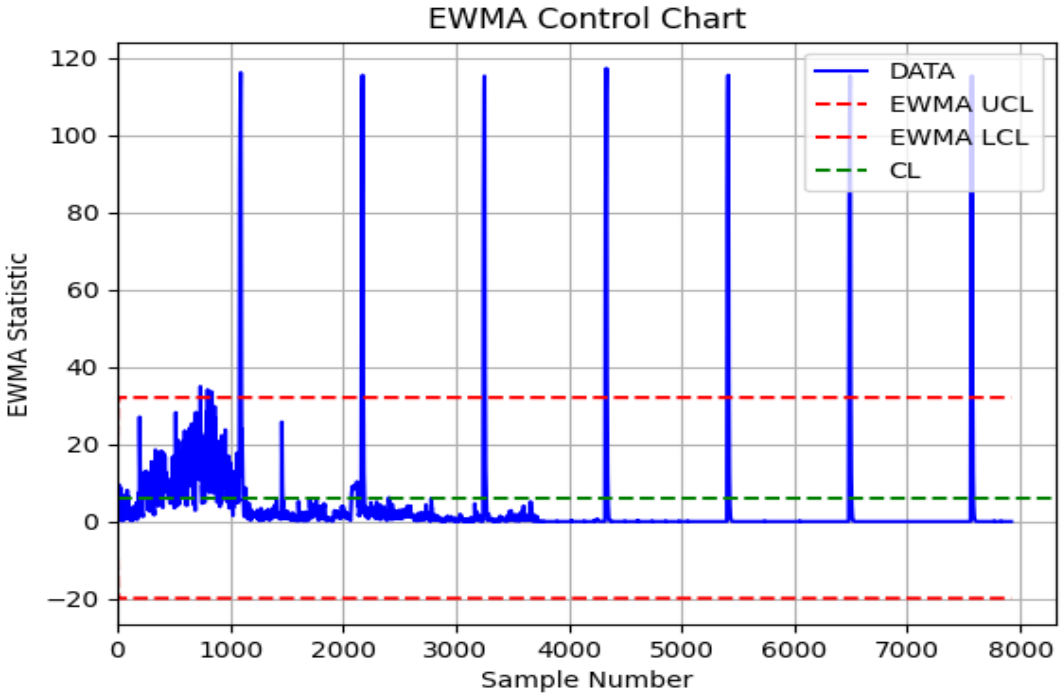


Figure III.25: EWMA Control Chart for low intensity of SYN segments

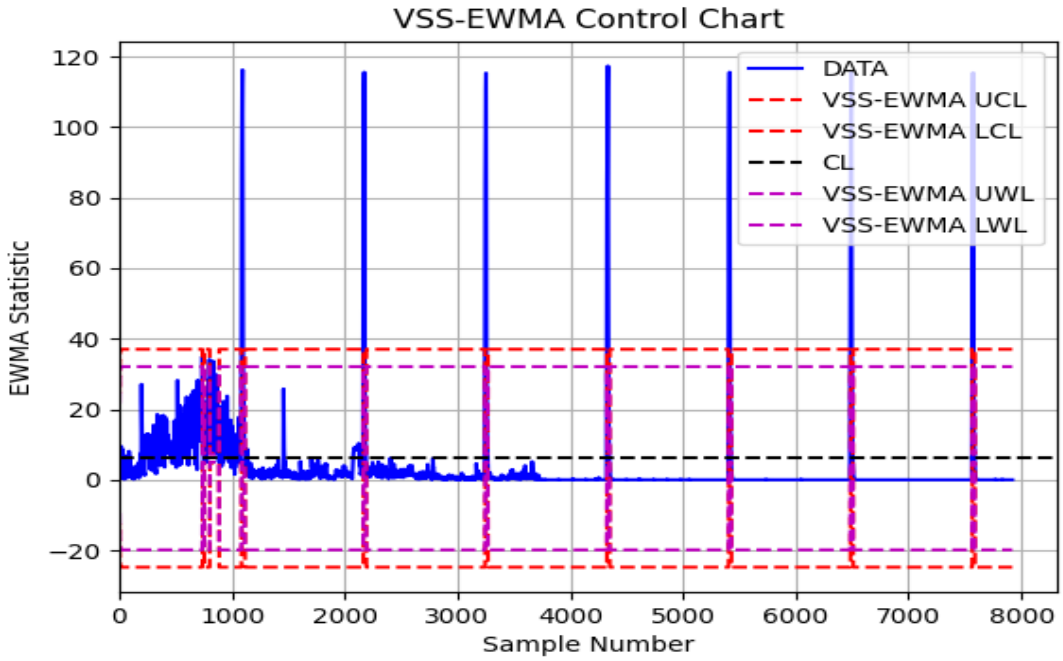


Figure III.26: VSS-EWMA Control Chart for low intensity of SYN segments

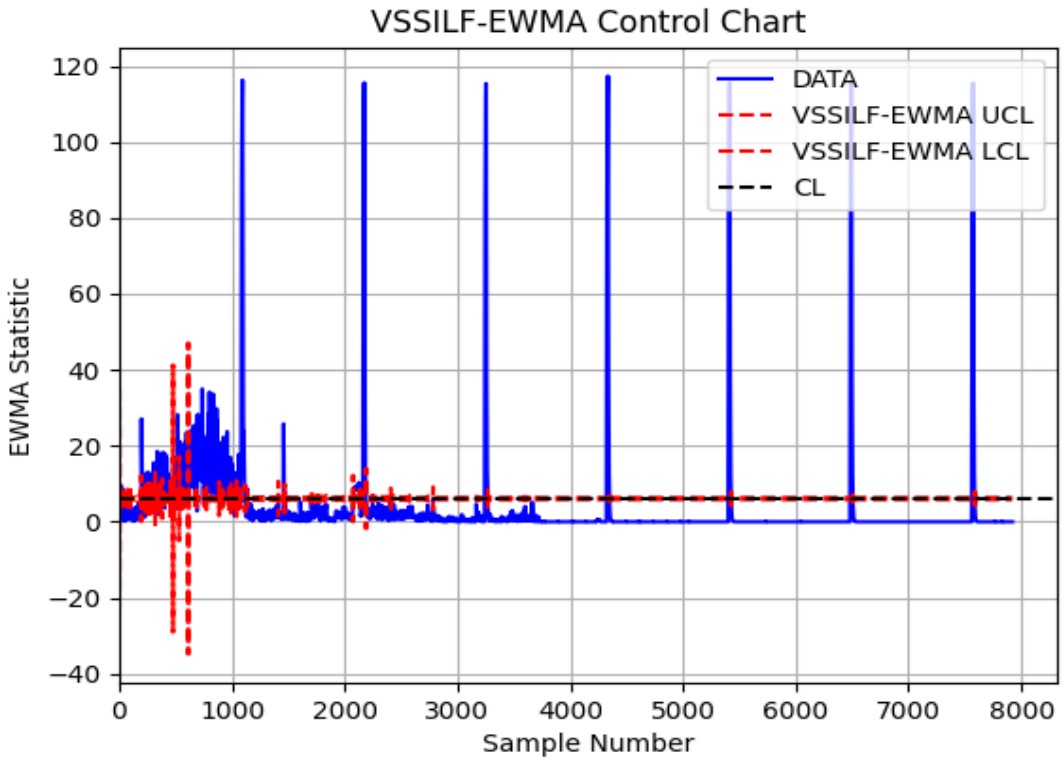


Figure III.27: VSSILF-EWMA Control Chart for low intensity of SYN segments

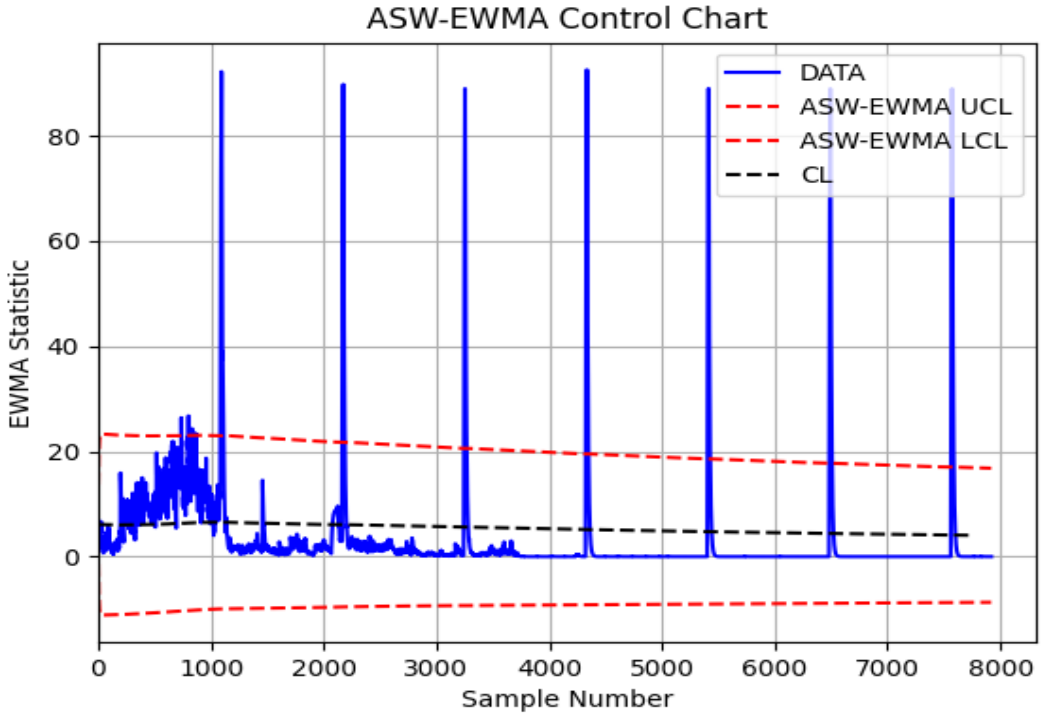


Figure III.28: ASW-EWMA Control Chart for low intensity of SYN segments

III.6.3 Smurf Attack Detection

To examine the effectiveness of Adaptive EWMA control chart models (including VSS-EWMA, VSSILF-EWMA, and ASW-EWMA Control Chart) in detecting Smurf attacks, we'll evaluate their performance across three distinct scenarios: random traffic featuring ICMP Echo Reply message flows, traffic subjected to high-intensity Smurf attacks, and traffic experiencing low-intensity Smurf attacks. Our analysis will utilize training data sourced from the fifth day of the second week and testing data extracted from the first day of the fifth week of the dataset.

III.6.4 The result of Smurf attack traffic

III.6.4.1 Random traffic of ICMP Echo Replay flow

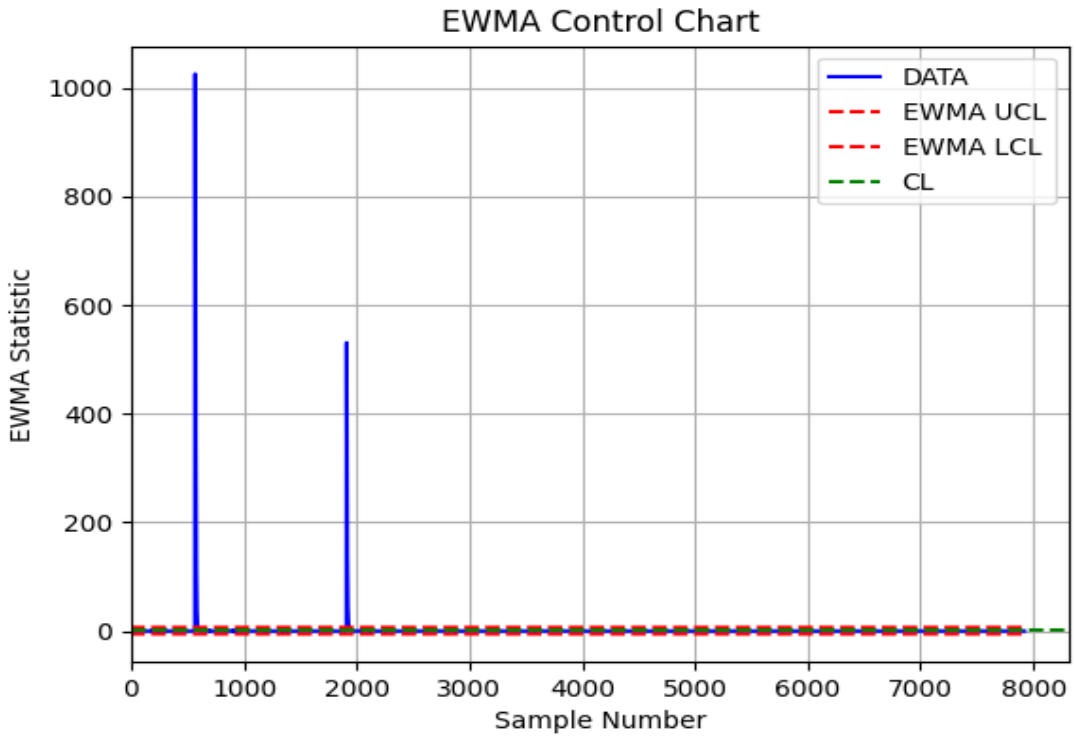


Figure III.29: EWMA Control Chart for the flow of echo replay message

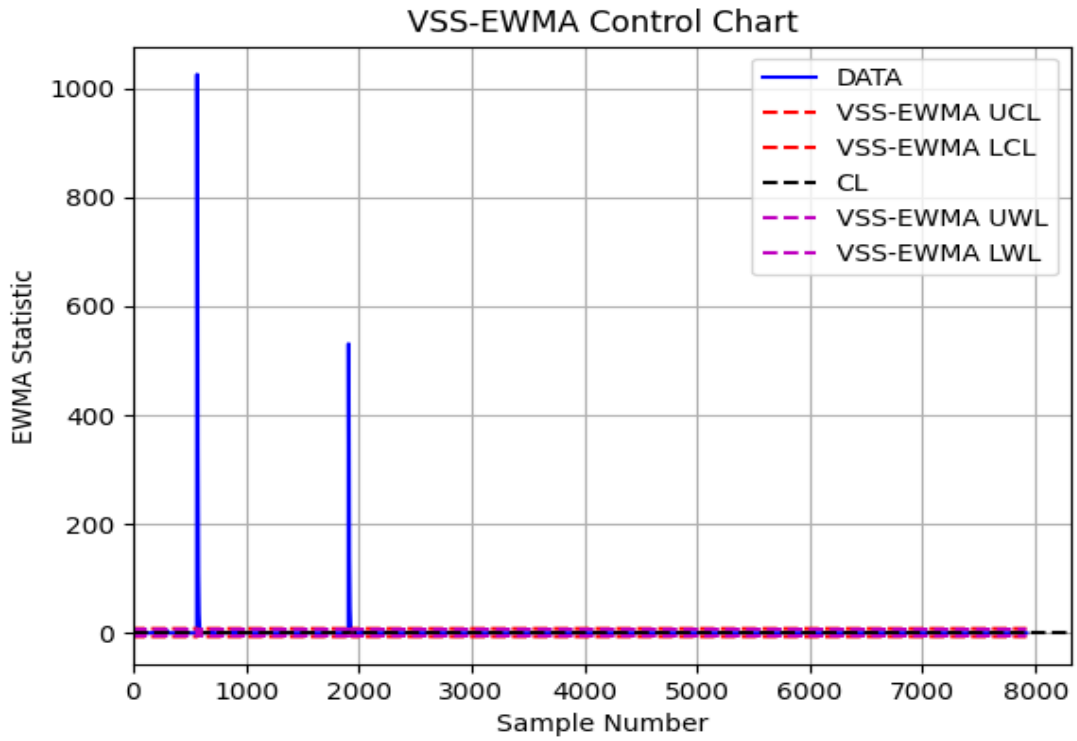


Figure III.30: VSS-EWMA Control Chart for the flow of echo replay message

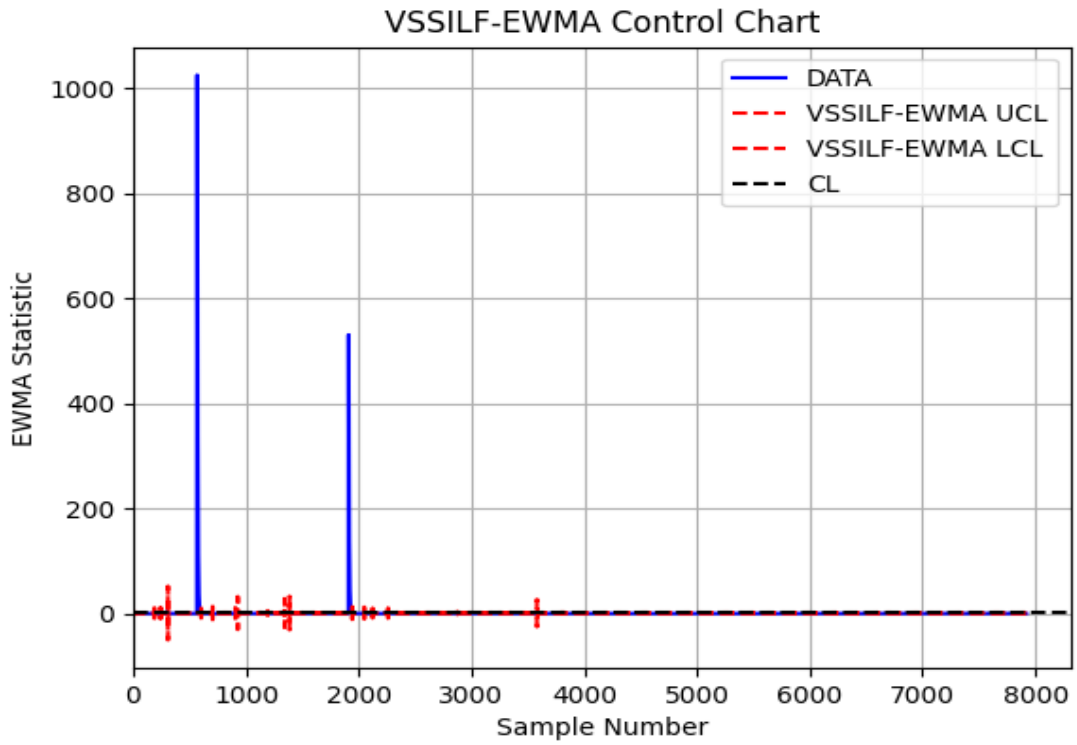


Figure III.31: VSSILF-EWMA Control Chart for the flow of echo replay message

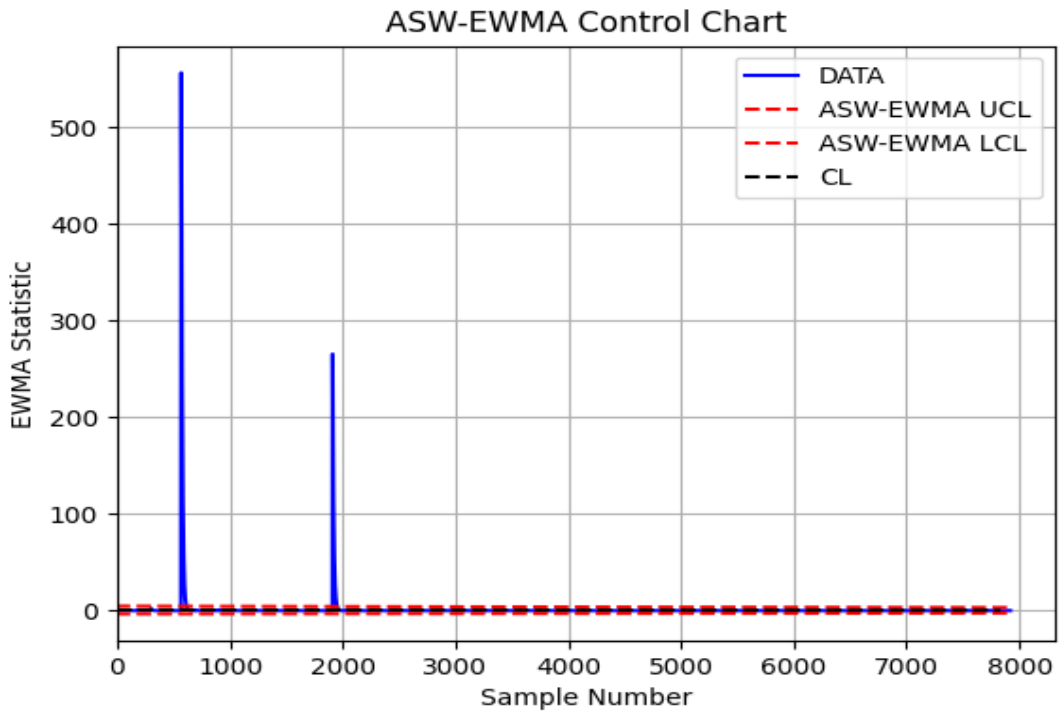


Figure III.32: ASW-EWMA Control Chart for the flow of echo replay message

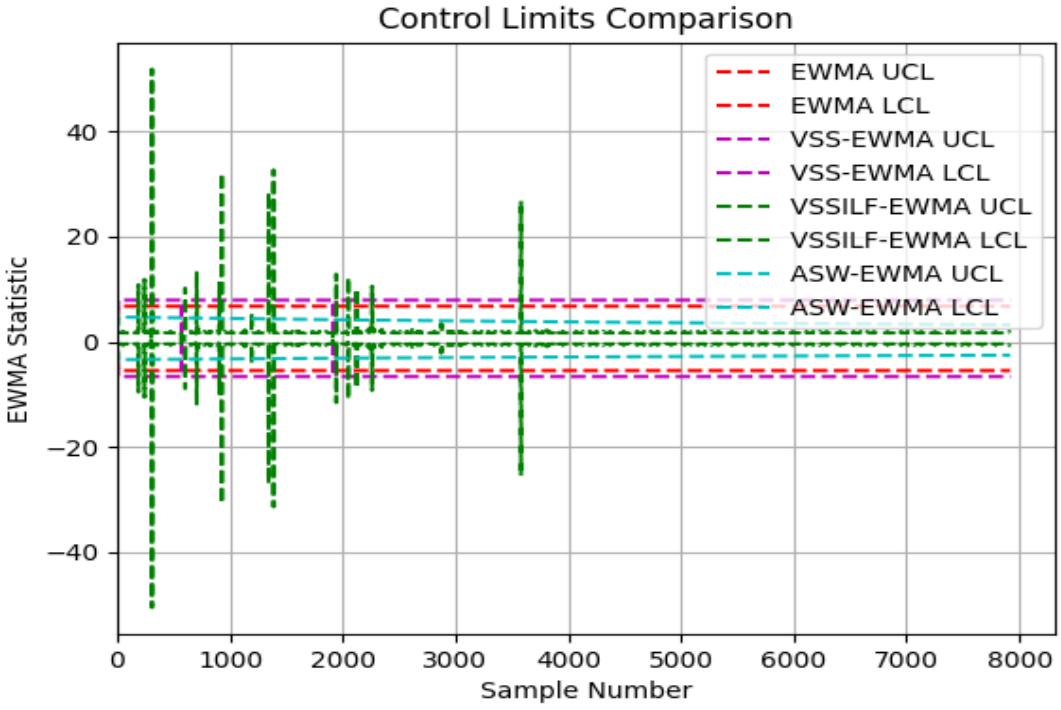


Figure III.33: Comparison of control Charts limits for the flow of echo replay message

III.6.4.2 Traffic with high intensity Smurf attacks

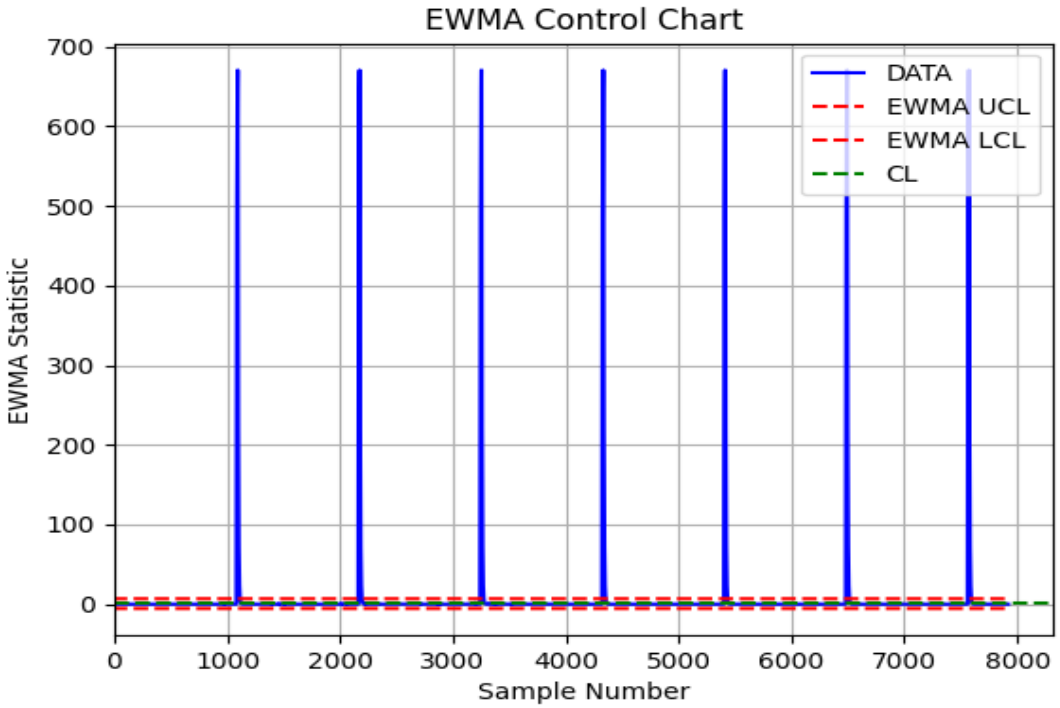


Figure III.34: EWMA control chart for high intensity Smurf attacks

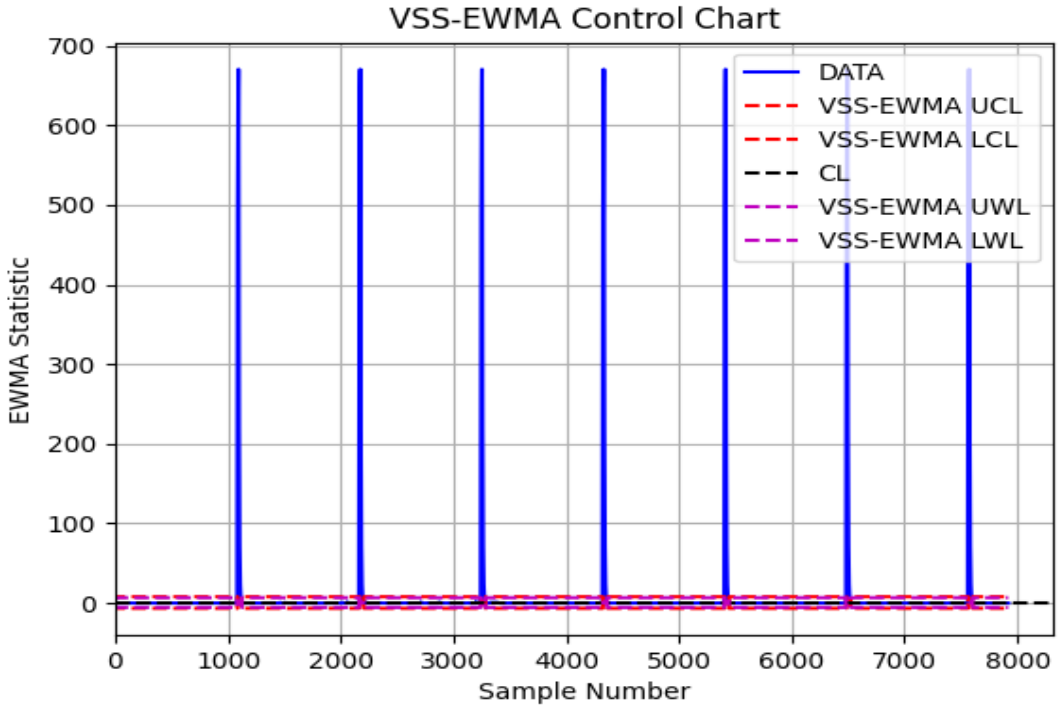


Figure III.35: VSS-EWMA control chart for high intensity Smurf attacks

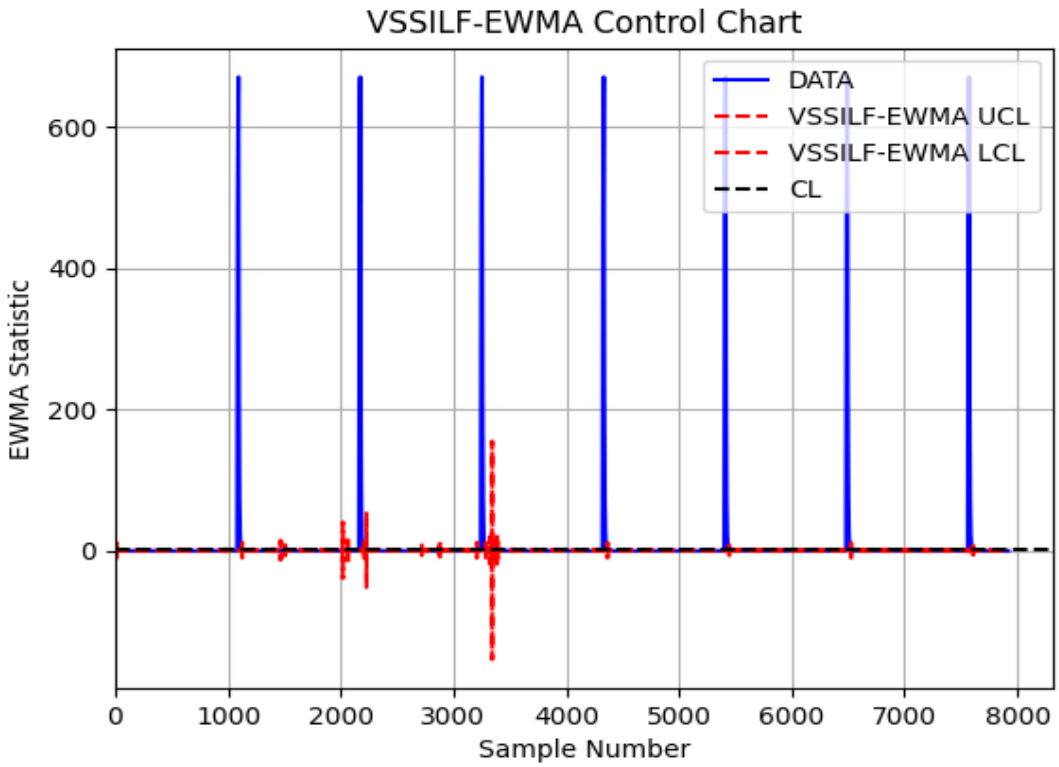


Figure III.36: VSSILF-EWMA control chart for high intensity Smurf attacks

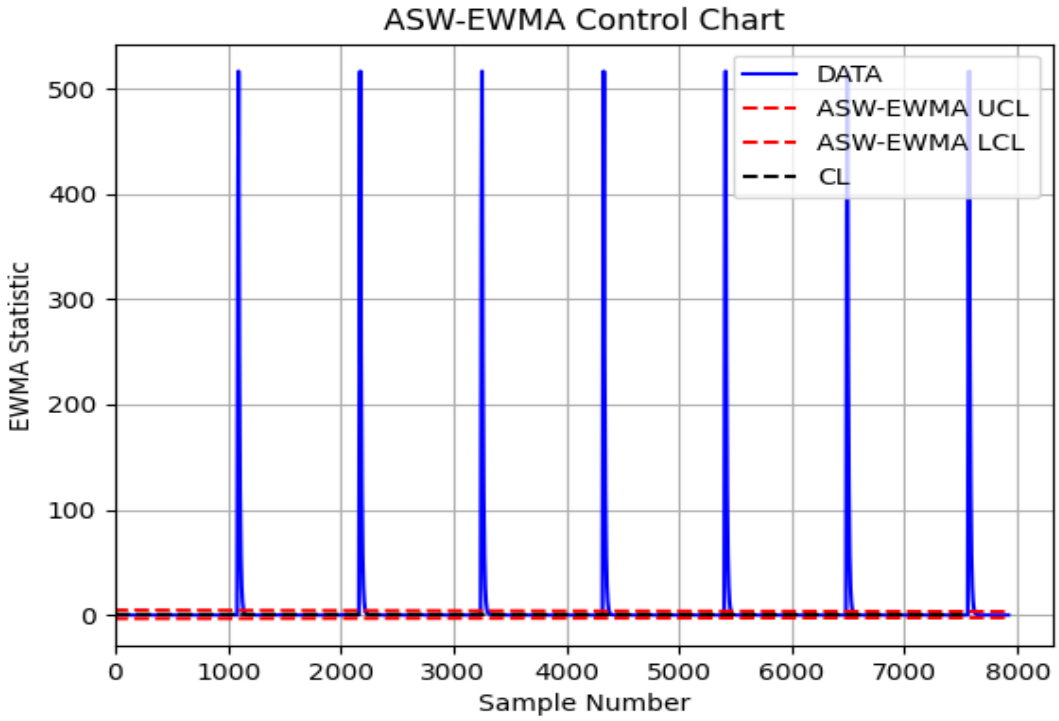


Figure III.37: ASW-EWMA control chart for high intensity Smurf attacks

III.6.4.3 Traffic with low intensity Smurf attacks

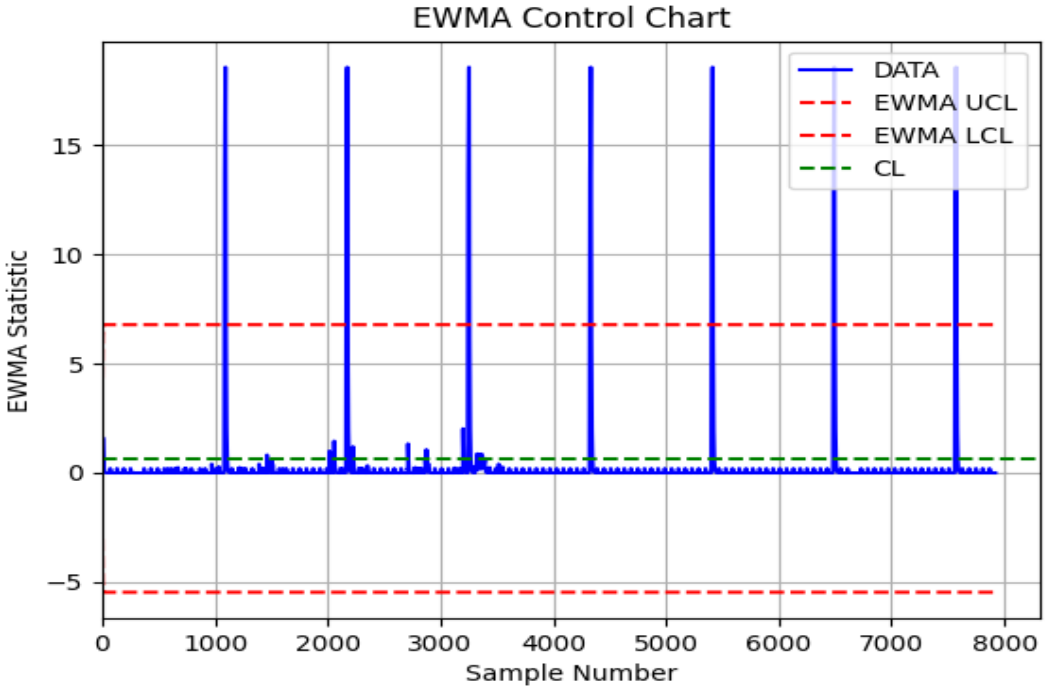


Figure III.38: EWMA control chart for low intensity Smurf attacks

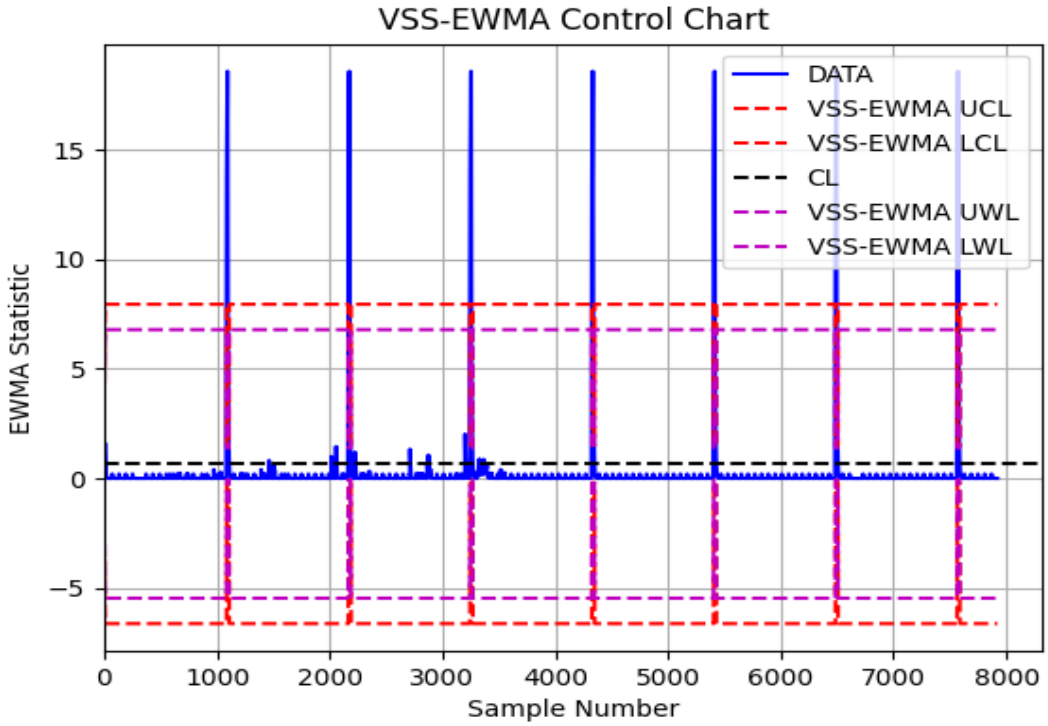


Figure III.39: VSS-EWMA control chart for low intensity Smurf attacks

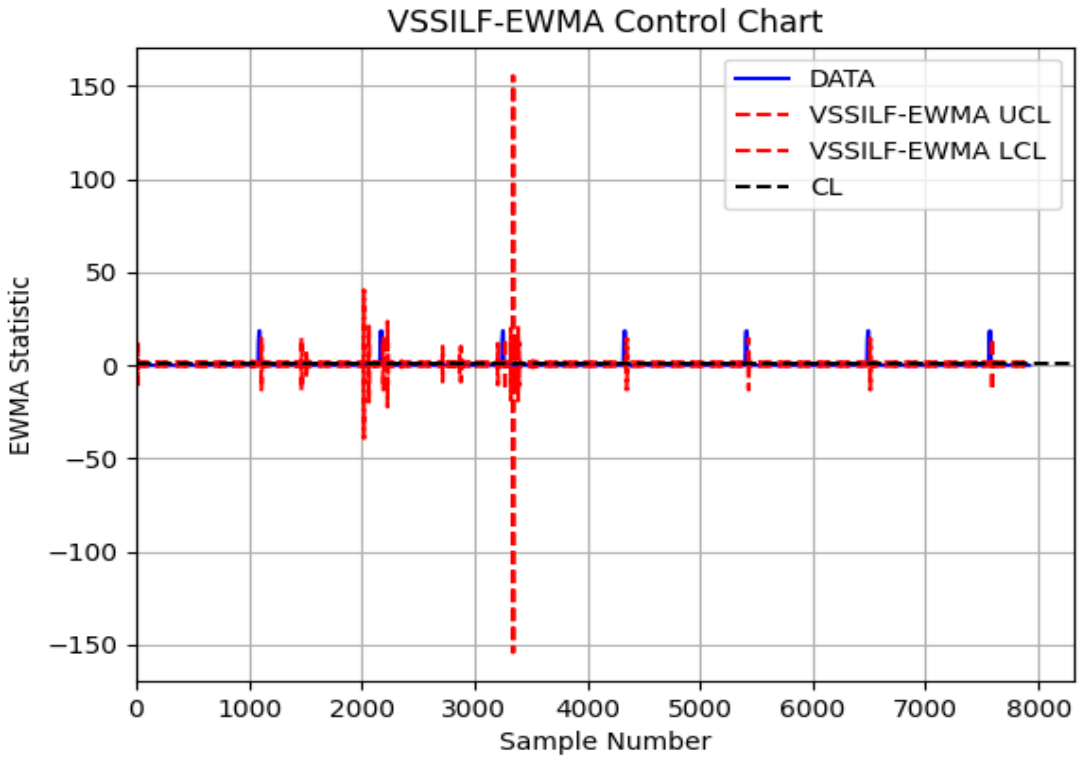


Figure III.40: VSSILF-EWMA control chart for low intensity Smurf attacks

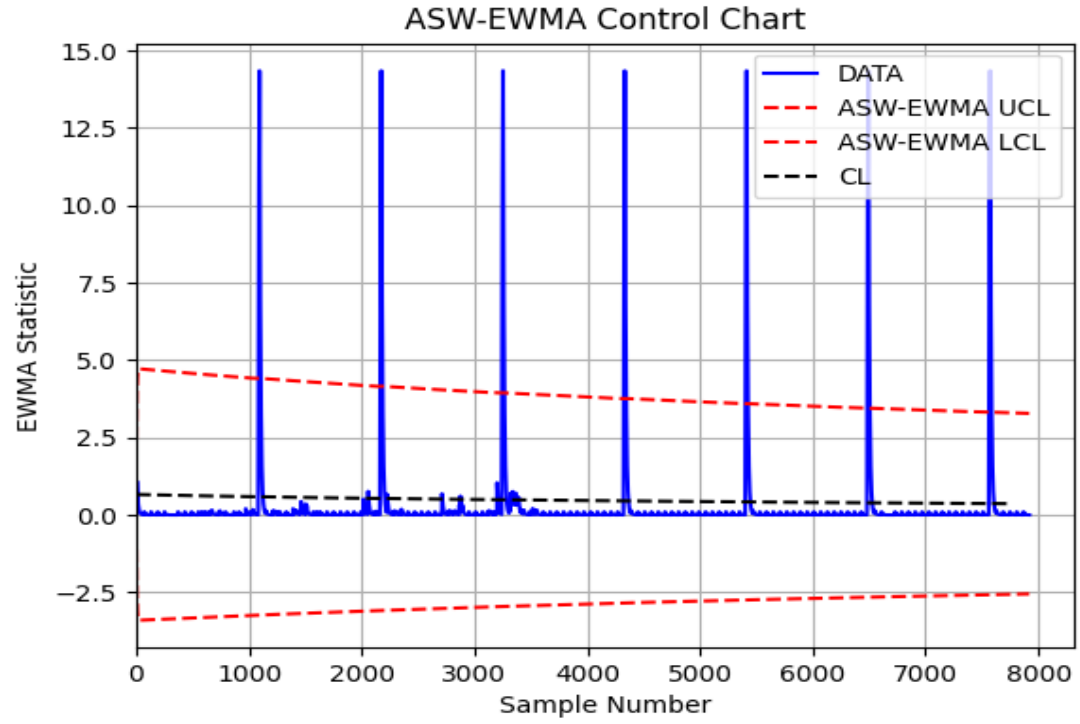


Figure III.41: Adaptive ASW-EWMA control chart for low intensity Smurf attacks

III.7. Results and discussions

In our simulation, we examined the impact of varying the smoothing parameter (**λ**) and the width of the control limits (**L**) on control chart limits using SYN and Smurf attack scenarios with random data. By fixing **λ** and altering **L**, we observed that larger values of **L** resulted in wider control limits. Conversely, when L was held constant and λ was varied, increasing **λ** also led to wider control limits.

We utilized EWMA control charts along with several Adaptive EWMA control chart models specifically VSS-EWMA, VSSILF-EWMA, and ASW-EWMA to detect DoS and DDoS attacks, including TCP SYN Flood and Smurf attacks, using the DARPA99 dataset. Our analysis revealed that these attacks caused a significant increase in control chart statistics compared to normal conditions. During normal operation, the control chart statistics remained low, but they increased markedly during an attack.

The EWMA control chart effectively detected all the attacks, though it occasionally generated false alarms by indicating normal traffic as anomalous. The VSS-EWMA charts also detected all attacks with some alerts. A key advantage of the VSS-EWMA charts is their adaptive control limits, which narrow significantly before a process exhibits high variability, thereby enhancing sensitivity to sudden changes. The VSSILF-EWMA charts, on the other hand, had very narrow control limits, causing them to classify a large portion of the data as attacks. This resulted in a high number of false alarms, reducing their practical usability in distinguishing between normal and malicious traffic. The ASW-EWMA detected attacks with similar effectiveness to the standard EWMA control chart. However, the integration Sliding Window provided an additional layer of adaptability, potentially improving detection accuracy and reducing false alarms over time.

Overall, all control chart models successfully identified attacks in traffic containing TCP SYN segments and ICMP Echo Reply messages or in high-intensity traffic. By adjusting the values of **λ** and **L**, it is possible to reduce the number of false alerts, highlighting the importance of fine-tuning these parameters based on specific network traffic characteristics and desired sensitivity levels.

III.8. Conclusion

Based on the simulation analysis of using Adaptive EWMA control charts to detect DoS and DDoS attacks (TCP SYN Flood and Smurf) from the DARPA99 dataset, we conclude that control charts are effective for detecting such attacks. All four control charts (standard EWMA, VSS-EWMA, VSSILF-EWMA, and ASW-EWMA) were able to detect attacks in high-intensity traffic or traffic containing TCP SYN segments and ICMP Echo Reply messages. However, some false alarms occurred, with the VSSILF EWMA chart being particularly prone to a high number of false positives due to its narrow control limits. The adaptability of the VSS-EWMA and ASW-EWMA models provided a balance between sensitivity and false alarm rates, demonstrating the importance of fine-tuning the parameters λ and **L** to optimize detection performance. Overall, the study highlights the effectiveness of these control charts in safeguarding network integrity against diverse attack scenarios.

General Conclusion

General Conclusion

The research conducted represents a significant advancement in understanding the efficacy of EWMA control charts and adaptive control charts in detecting and mitigating DoS and DDoS attacks within network traffic. In today's interconnected world, where networks serve as the backbone of modern communication and commerce, the threat of cyber-attacks looms large. DoS and DDoS attacks, in particular, pose substantial risks, capable of causing widespread disruption, financial loss, and reputational damage to targeted entities.

Amidst this landscape of evolving threats, the need for robust detection mechanisms becomes imperative. EWMA control charts, renowned for their sensitivity to subtle changes in process behavior, offer a promising avenue for anomaly detection in network traffic. By monitoring exponentially weighted moving averages of observations, EWMA charts provide a real-time assessment of deviations from expected patterns, enabling timely intervention in the event of an attack.

Furthermore, the integration of adaptive control chart models, such as VSS-EWMA, VSSILF-EWMA, and ASW-EWMA, represents a significant leap forward in detection capabilities. These adaptive models exhibit dynamic responsiveness to changing network conditions, adjusting control limits and thresholds in real-time to optimize detection sensitivity while minimizing false alarms. VSS-EWMA charts, for instance, demonstrate adaptability through the narrowing of control limits in response to heightened variability, thereby enhancing sensitivity to subtle attack patterns.

The research findings, derived from meticulous simulations utilizing the DARPA99 dataset, offer valuable insights into the performance of these control chart models under diverse attack scenarios. While EWMA control charts exhibit commendable effectiveness in detecting attacks, adaptive models showcase superior sensitivity and resilience in the face of evolving threats. ASW-EWMA, leveraging machine learning techniques, stands out for its ability to dynamically adapt to emerging attack patterns, offering a proactive defense against cyber threats.

However, it is essential to acknowledge the inherent challenges in assuming normality in network traffic distribution, as deviations from this assumption can lead to erroneous detection outcomes. Addressing this limitation requires ongoing research efforts aimed at

refining detection methodologies to better align with the complex realities of network traffic behavior.

Finally, this research underscores the pivotal role of EWMA control charts and adaptive control charts in fortifying network security against cyber threats. By harnessing the capabilities of these models and continually refining detection strategies, organizations can bolster their resilience and effectively safeguard network infrastructures from the ever-evolving threat landscape, ensuring the uninterrupted flow of communication and commerce in our interconnected world.

Bibliography

Bibliography

[1] Bonaventure, O. (2012). Computer Networking: Principles, Protocols and Practice. Publisher: Lulu.com.

[2] Kurose, J., & Ross, K. (2016). Computer Networking: A Top-Down Approach. Publisher: Pearson.

[3] Goralski, W. (2017). The Illustrated Network, Second Edition. Publisher: Morgan Kaufmann.

[4] Comer, D. E. (2006). INTERNETWORKING with TCP/IP principles, Protocols, And Architectures FOURTH EDITION. Publisher: Pearson.

[5] Forouzan, B. A. (2006). TCP/IP: Protocol Suite. Publisher: McGraw-Hill.

[6] Comer, D. (2017). Internetworking with TCP/IP - Principles, Protocols, and Architecture. Publisher: Pearson.

[7] Stevens, W. R. (1994). TCP/IP Illustrated, Volume 1: The Protocols. Publisher: Addison-Wesley.

[8] Stallings, W. (2016). Network Security Essentials (4th ed.). Publisher: Pearson.

[9] Erickson, J. (2008). Hacking: The Art of Exploitation (2nd ed.). Publisher: No Starch Press.

[10] Cybersecurity and Infrastructure Security Agency (CISA). (2020). Title. Publisher: U.S. Department of Homeland Security.

[11] Stallings, W. (2016). Cryptography and Network Security. Publisher: Pearson.

[12] Caravelli, J. (2015). Cyber Resilience: A New Paradigm for Cybersecurity. Publisher: Springer.

[13] Stuttard, D., & Pinto, M. (2021). Web Application Hacker's Handbook: Finding and Exploiting Security Flaws. Publisher: Wiley.

[14] IGI Global. (2021). Handbook of Research on Information and Cyber Security in the Fourth Industrial Revolution. Publisher: IGI Global.

[15] McNicholas, E., & Mohan, V. (2019). Cybersecurity: A Practical Guide to the Law of Cyber Risk. Publisher: LexisNexis.

[16] Mirkovic, J., Dietrich, S., Dittrich, D., & Reiher, P. (2010). DDoS Attacks: Evolution, Detection, Prevention, Reaction, and Tolerance. Publisher: Springer.

[17] Akerlof, G. A., & Shiller, R. J. (2015). Phishing for Phool's: The Economics of Manipulation and Deception. Publisher: Princeton University Press.

[18] Ligh, M. H., Adair, S., Hartstein, B., Richard, M., & Ligh, M. H. (2011). Malware Analyst's Cookbook and DVD: Tools and Techniques for Fighting Malicious Code. Publisher: Wiley.

[19] Mitnick, K. D. (2002). The Art of Deception: Controlling the Human Element of Security. Publisher: Wiley.

[20] Clarke, J. (2012). SQL Injection Attacks and Defense. Publisher: Syngress.

[21] Van Beijnum, I. (2002). BGP. Publisher: O'Reilly Media.

[22] Mitnick, K. D. (2005). The Art of Intrusion: The Real Stories Behind the Exploits of Hackers, Intruders and Deceivers. Publisher: Wiley

[23] BOUYEDDOO, B. (2021). Détection avancée des anomalies et applications aux systèmes de communication. Université Abou Bekr Belkaid - Tlemcen.

[24] Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product.

[25] Montgomery, D. C. (2009). Introduction to Statistical Quality Control. Hoboken, NJ: John Wiley & Sons.

[26] El-Maleh, A. H. A. (2011). Statistical process control: Theory and practice. LAP Lambert Academic Publishing.

[27] Levinson, W. A. (1958). Statistical Quality Control in World War II Production.

[28] Schilling, E. G. (1942). The Role of Statistics in War Production.

[29] Juran, J. M. (2010). Juran's Quality Control Handbook.

[30] Grant, E. L., & Leavenworth, R. S. (1980). Statistical Quality Control. New York, NY: McGraw-Hill.

[31] Wheeler, D. J. (1993). Understanding Variation: The Key to Managing Chaos. Knoxville, TN: SPC Press.

[32] Oh, A. (2024, March 14). Control Charts: Everything You Need to Know. ClearPoint Strategy. Retrieved from [https://www.clearpointstrategy.com/control-charts/]

[33] Achieve Process Excellence. (2024, March 7). The Ultimate Guide to Control Charts in Six Sigma. Retrieved from [https://www.6sigma.us/process-improvement/control-charts-six-sigma-ultimate-guide/]

[34] Suman, G., & Prajapati, D. (2018). Control chart applications in healthcare: a literature review. International Journal of Metrology and Quality Engineering, 9(4), 401-412.

[35] Montgomery, D. C. (2009). Statistical Quality Control. Hoboken, NJ: John Wiley & Sons.

[36] Montgomery, D. C., & Anderson-Cook, C. M. (2012). Introduction to Statistical Quality Control. Hoboken, NJ: John Wiley & Sons.

[37] Wheeler, D. J. (1995). Statistical Process Control: A Practical Guide. Knoxville, TN: SPC Press.

[38] Wheeler, D. J. (1991). Understanding Statistical Process Control. Knoxville, TN: SPC Press.

[39] Roberts, S. W. (1959). Control chart tests based on geometric moving averages. Technometrics, 1(3), 239-250.

[40] Montgomery, D. C. (2005). Introduction to Statistical Quality Control (5th ed.). Hoboken, NJ: John Wiley & Sons.

[41] Montgomery, D. C. (2013). Introduction to Statistical Quality Control (7th ed.). Hoboken, NJ: John Wiley & Sons.

[42] Benneyan, J. C., Lloyd, R. C., & Plsek, P. E. (2003). Statistical process control as a tool for research and healthcare improvement. Quality and Safety in Health Care, 12(6), 458-464.

[43] Box, G. E. P., Jenkins, G. M., & Gregory, C. (2008). Time Series Analysis: Forecasting and Control. Hoboken, NJ: John Wiley & Sons.

[44] Breyfogle III, F. W. (2003). Implementing Six Sigma: Smarter Solutions Using Statistical Methods. Hoboken, NJ: John Wiley & Sons.

[45] Adeoti, O. A. (2018). A new double exponentially weighted moving average control chart using repetitive sampling. International Journal of Quality & Reliability Management, 35(7), 1419-1432.

[46] Wang, J., Arslan, M., Riaz, A., et al. (2023). Triple exponentially weighted moving average control chart with measurement error. Scientific Reports, 13, 14760.

[47] Letshedi, T. I., Malela-Majika, J.-C., Castagliola, P., & Shongwe, S. C. (2021). Distribution-free triple EWMA control chart for monitoring the process location using the Wilcoxon rank-sum statistic with fast initial response feature. Quality and Reliability Engineering International, 37(5), 1996-2013.

[48] Costa, A. F. B., & De Magalhães, M. S. (2007). Adaptive exponentially weighted moving average control charts. Journal of Statistical Planning and Inference, 137(7), 2182-2197.

[49] Amiri, A., Nedaie, A., & Alikhani, M. (Year). A New Adaptive Variable Sample Size Approach in EWMA Control Chart. Industrial Engineering Department, Shahed University, Tehran, Iran.

[50] De Magalhães, M. S., et al. (2009). Statistical design of a hierarchy of two states adaptive parameters EWMA control charts. European Journal of Operational Research, 199(1), 68-82