I. J. Computer Network and Information Security, **2021, 6, 13-25**
Published Online December 2021 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2021.06.02

A Comprehensive Review of Intrusion Detection and Prevention Systems against Single Flood Attacks in SIP-Based Systems

Sheeba. Armoogum
University of Mauritius / Faculty of Information, Communication and Digital Technologies, Reduit, Mauritius
E-mail: s.armoogum@uom.ac.mu

Nawaz. Mohamudally
University of Technology Mauritius / School of Innovative Technologies and Engineering, La Tour Koenig, Mauritius
E-mail: alimohamudally@umail.utm.ac.mu

Received: 20 April 2021; Revised: 21 May 2021; Accepted: 28 June 2021; Published: 08 December 2021

Abstract: Voice over Internet Protocol (VoIP) is a recent voice communication technology and due to its variety of calling capabilities, the system is expected to fuel the market value even further in the next five years. However, there are serious concerns since VoIP systems are frequently been attacked. According to recent security alliance reports, malicious activities have increased largely during the current pandemic against VoIP and other vulnerable networks. This hence implies that existing models are not sufficiently reliable since most of them do not have a hundred percent detection rate. In this paper, a review of our most recent Intrusion Detection & Prevention Systems (IDPS) developed is proposed together with a comparative analysis. The final work consisted of ten models which addressed flood intentional attacks to mitigate VoIP attacks. The methodological approaches of the studies included the quantitative and scientific paradigms, for which several instruments (comparative analysis and experiments) were used. Six prevention models were developed using three sorting methods combined with either a modified galloping algorithm or an extended quadratic algorithm. The seventh IDPS was designed by improving an existing genetic algorithm (e-GAP) and the eighth model is a novel deep learning method known as the Closest Adjacent Neighbour (CAN). Finally, for a better comparative analysis of AI-based algorithms, a Deep Analysis of the Intruder Tracing (DAIT) model using a bottom-up approach was developed to address the issues of processing time, effectiveness, and efficiency which were challenges when addressing very large datasets of incoming messages. This novel method prevented intruders to access a system without authorization and avoided any anomaly filtering at the firewall with a minimum processing time. Results revealed that the DAIT and the e-GAP models are very efficient and gave better results when benchmarking with models. These two models obtained an F-score of 98.83%, a detection rate of 100%, a false rate of 0%, an accuracy of 98.7%, and finally a processing time per message of 0.092 ms and 0.094 ms respectively. When comparing with previous models in the literature from which it is specified that detection rates obtained are 95.5% and false-positive alarm of around 1.8%, except for one recent machine learning-based model having a detection rate of 100% and a processing time of 0.53 ms, the DAIT and the e-GAP models give better results.

Index Terms: Voice over Internet Protocol (VoIP), Denial of Service (DoS), Flood Attacks, Session Initial Protocol (SIP), Intrusion Detection and Prevention System (IDPS), Deep Analysis, Deep Learning, Genetic Algorithm.

1. Introduction

In the last ten years, the world has witnessed a growth in the Voice over Internet Protocol (VoIP) market thanks to the effort taken by governments, regulatory bodies, service providers and customers. The rise in market demand is also due to the growing trend of workforce mobility. Global Market Insights (a worldwide market consultant company) points out that by deploying VoIP solutions, enterprises can improve communication and collaborations among employees and remote users to increase business productivity [1]. With its variety of calling capabilities, VoIP is expected to fuel the VoIP market value even further in the next five years. According to Chauhan et al. [2], the average rate of increase in VoIP residential subscriptions per year is 7.58%, which is depicted in Fig.1 (a projection until the year 2025).

The above benefits indicate that VoIP technology contains various features that are attracting businesses and customers. The technical reason is that this technology has a set of protocols that allow multimedia communication over the same IP network that is used for the transfer of data and files (whether wireless or wired) as well as provides several

possibilities for migration from a PSTN network. As mentioned by Chauhan et al. [2] and many other researchers (Ehlert et al. [3], Azad et al. [4] and Nazih et al. [5]), the Session Initiation Protocol (SIP) is the most popular VoIP protocol that is used for signaling multimedia calls over IP. Using this protocol, the signal during the establishment, execution, and termination phases of multimedia sessions between two or more participants can be easily been controlled.

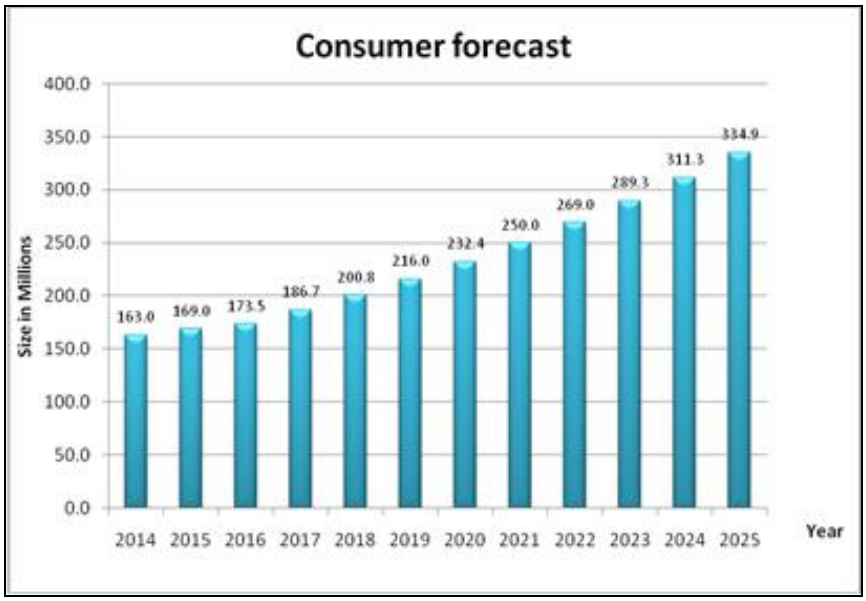

Fig.1. VoIP Residential Subscribers (2014-2025).

Having said that, despite the numerous research work that has been carried out to secure VoIP networks, official reports (IBM, Cisco, VOIPSA) put in the picture that this does not discourage hackers from conducting malicious activities on private and public networks. The core concern of this current study is to conduct literature on Flood-based Denials of Service (DoS) attacks for the reason that they are been treated as a major threat to the VoIP environment. In fact, these attacks are been executed by hackers from remote locations using an immense volume of fake traffic to occupy the resources to forbid legitimate traffic to pass through. Mostly, the attackers' target devices are the server and vulnerable clients in private networks, forcing them to consume all of their resources (processor, memory and bandwidth). In such cases, either the server/clients may crash or a reboot of the system may be required. The situation is worst in case a network has a DDoS attack, which normally occurs when multiple attackers flood the resources of the targeted server/clients. DDoS attacks represent the central threat observed by VoIP providers. Equipped with free numerous tools, attackers carry out DDoS attacks to exploit vulnerabilities in the network. Cisco [6] reported that there has been a global loss of $150 million in 2019. There are still gaps in the research regarding the effectiveness, efficacy and security measures. For instance, it was observed that, though the researchers used different techniques to mitigate the attacks, there was a lack of benchmarking of their results with other works. Most of the studies have mitigation values of between 75% to less than 95.5% [7], while for others, the proxy crashed at a very small shared memory in a small duration of time.

The main significance of the study is to mitigate the attacks to VoIP systems, on people and data so as to find solutions for critical situations like the recently triggered pandemic where people are largely in the use of VoIP systems. This leads to the importance of designing models to enhance the existing security systems and evaluating them to help manufacturers to improve their IDPS.

Recently, the authors of the current research study have developed nine IDPS to address the issues of flood attacks. These models were developed using conventional algorithms, artificial intelligence and deep learning algorithms. The purpose of this paper is to conduct a comprehensive assessment of the nine flood mitigation models for VoIP systems to determine the efficacy and efficiency, taking account of several key performance indicators like the detection rate, F-measure, memory consumption, CPU utilization amongst others. The specific objective is to determine the most reliable models which can be integrated into a defense security system that can address Denial of Service flood attacks whilst benchmarking them with models developed by the research community and industry.

This paper is planned as follows: Section 2 covers a list of related work and a précis of the nine models. The research methodology used to develop the models is explained in Section 3. Section 4 discusses the deep analysis among the models while an in-depth discussion is presented in Section 5. Finally, a conclusion is explained in the last section.

2. Background Study

This section summarises the different efforts conducted by the research community on flood attacks to fight against intruders and also an attempt to mitigate network issues. A small brief of each model developed by the authors of this current study is also presented.

2..1. Related Work

Prior to giving a brief about the models developed, a survey of other prevention tools and models on different attacks on VoIP systems is discussed in this section.

Currently, many VOIP projects are been deployed around the world because of their easy way of installation [1]. Discussions were made at different platforms (conferences, workshops, summits) and yet the security issue is still a challenge for VOIP providers. To eliminate attacks against VoIP systems is almost impossible despite the many efforts emanated by industry and researchers.

Many flood-based mechanisms were developed since the inception of VoIP technology. The tool proposed by Chen [8] addressed attacks by modifying original-state machines to detect fraudulent activities in systems. A novel model was proposed by Li et al. [9] to detect DoS flood attacks to tackle the issue of high attack rates and to reduce the processing time. The model achieved successful results using a method known as the sliding window to achieve a low false rate for a large amount of traffic. Similarly, Wan et al. [10] proposed an efficient method using a queue analysis model to reduce CPU utilization.

A genetic algorithm-based mitigation technique was developed by Dhak & Lade [11] in an attempt to reduce the high false alarm rate. The model is able to give positive results using a set of defined rules using the support of a firewall. In the same line, a signature and anomaly-based model was developed by Ahmed & Ali [12] to address various types of flood behaviours. The authors' reports recently have been considered as references and based on their investigations, a threshold values of 95.5% for the detection rate and a false positive are of 1.8% are to be considered as benchmarks. Similarly, Akbar and Muddassar [13] proposed a mitigation technique to address composite attacks on IP telephony networks. Their results revealed that a detection rate of above 98% could be achieved. Tang et al [14] proposed a Hellinger distance method to combat flood attacks. When benchmarking with other models, the authors in [12] mentioned that their model gave better accuracy.

2.2. Synopsis of the Nine Prevention Models

To address the issues of DoS flood attacks, the following models were developed:

1) Three extended hybrid algorithms using Galloping search with Merge sort, Quicksort, and Heap sort algorithms.
2) Three extended hybrid algorithms using Quadratic search with Merge sort, Quicksort, and Heap sort algorithms.
3) An Extended Genetic Algorithmic approach.
4) A Deep Learning algorithm using a novel Closest Adjacent Method
5) A new Deep Analysis Technique using the Dictionary method.

Sorted Galloping and Quadratic Prevention Methods: These two groups of defense mechanisms were devised by using three types of sorting algorithms, combining with either a modified galloping algorithm or an extended quadratic algorithm, and are known as the Sorted Galloping Prevention (SGP) [15] and the Sorted Quadratic Prevention (SQP) [16] algorithms respectively. The galloping-based and quadratic-based prevention methods were combined to give rise to six different prevention approaches (m-SGP, q-SGP, h-SGP, m-SQP, q-SQP, and h-SQP).

The input to the system is a CSV file in an ordered time field from the firewall. The CSV file is been cleansed on the four fields Source, Destination, Time and Info. A two-dimensional list is created consisting of the source address, destination address and Info fields which makes the statistical algorithms different from other work developed by previous researchers. The source list is been further sorted using either a merge sort, a quick sort, or a heap sort algorithm on the source address field [15, 16]. The next phase of the algorithm is to loop the entire list to search for the target, which is been assigned as the first element of the list for each iteration based on few conditions. If the searched target is been repeated more than once with the same Info within an interval of time of 500 ms, or in case the Info field is different for the same Source field of the target for the same period, then the targeted element is quarantined and is sent to another list. Otherwise, the target is added to another list (legitimate list) [15, 16]. Upon completion of the loop, the quarantine list is further analyzed to detect the targeted users of attacks and to inform the system administrator via a monitoring and analysis station accordingly. The unique legitimate source IP addresses are then forwarded to the firewall.

Extended Genetic Algorithm Prevention Model: The defense mechanism, known as Extended Genetic Algorithm Prevention (e-GAP), was designed by improving an existing genetic algorithm used in biological areas [17]. Genetic algorithms are been used in diversified fields to find solutions for several optimization problems mostly due to its

natural selection of the fittest individual for the reproduction of offspring of the next generation based on Charles Darwin's theory of evolution [18]. In the concept of genetic algorithms, the process starts with an initial population and repeatedly creates a new population until an ultimate solution is determined. Since genetic algorithms are conceptualized on random selection, it will be difficult to identify the exact number of generations. A genetic algorithm differs from other algorithms in terms of the fittest variable concept with crossover and mutation methods. A crossover method creates an offspring randomly from selected parents based on the fitness value and, then further mutates to create a new population. The outcome will be the individuals having the highest fitness value.

This evolutionary approach using a genetic algorithm technique was used previously to develop an IDS which can address the problem of a high false alarm rate triggered by illegitimate attackers and handlers [11]. The model captures a pool of messages and performs a cross-over process using the divide and swap concept to form two datasets. The latter undergoes a mutation process to identify the fitness value, to identify the legitimate and genuine messages based on a specific condition. The result of the mutation process is a new pool of messages and the process is repeated until no further message (individual) is been left over.

Closest Adjacent Neighbour Model: This new method, known as the Closest Adjacent Neighbour (CAN) algorithm [19] is designed, inspired by the K-Nearest Neighbours (KNN) algorithm and the Recurrent Neural Network (RNN) deep learning approach.

The CAN model identifies the next adjacent look-alike illegitimate messages in a given bulk of messages injected in a system and derives different patterns initiated by attackers with a minimum CPU utilization [19]. The rationale behind this proposed method is to identify the different patterns originated by attackers and classify them by the behaviour with minimal processing time. The innovative part of the model lies in the different aspects of detection and elimination of attack messages. Indeed, it is an architecture, starting with data cleansing at the start before converting the data into the recommended IDPS format. The model uses an internal sorting construct to search for similar targets and eliminate them based on the rules set by the algorithms [19]. The algorithm consists of several conditions intending to identify the various patterns emanated by the attackers and hence eliminate all non-genuine packets from the dataset. Next, the algorithm deals with issues related to false positives and false negatives to forbid flood-based attackers to trick the algorithms. The ultimate information obtained is sent to the firewall for further action.

Deep Analysis of Intruder Tracing (DAIT) using Bottom-Up Approach: This is another efficient proposal to conduct an accurate search for illegitimate messages that are hidden within a list of genuine packets using a deep learning concept and analyzed using a bottom-up approach [20]. This method is more efficient than the above-mentioned models as it can capture illegitimate messages from an unsorted list. The model first verifies the incoming list with the existing list of quarantined messages before cleansing the data. The model creates a two-dimensional list which is been created using the fields Source, Destination, and Info to identify unique source IP addresses from the list. These suspected unique addresses are sent to the firewall for further filtering following which they are sent to the administrator for further actions [20]. Finally, a deep analysis is conducted to detect any false positives and false negatives problems.

3. Review of Methods

The section covers the research methods used for preparing the data, designing the models and collecting the data using a testbed.

3.1. Research Design

Table 1 provides a summary of the different research types, methods, designs, and techniques used during the development of the nine models. Studies (Study 1 to Study 5) are constructive and empirical research work to analyze the security level of existing VoIP systems using new techniques. The research questions were addressed scientifically using the quantitative method.

The various mechanisms are summarised as follows:

1) Research Study 1: In this study, a mechanism was developed using three popular sorting algorithms and a search algorithm to prevent flood attacks. The aim was to develop three intrusion and mitigation systems using the Merge, the Quick, and the Heap sort algorithms using the Exponential (Galloping) search and compare them.
2) Research Study 2: The third study is similar to Study 1; however, in this case, to mitigate flood attacks, a quadratic search was modified. The aim was to compare the three mitigation algorithms developed.
3) Research Study 3: In this study, the IDPS was developed using a modified genetic algorithm to mitigate flood attacks.
4) Research Study 4: This was an attempt to mitigate flood attacks using a novel deep learning neighborhood technique.
5) Research Study 5: This was another novel deep analysis method to mitigate flood attacks known as deep analysis of intruder tracing using the bottom-up approach.

Table 1. Research Types, Methods, Design and Techniques

No	Study	Research Types	Research Methods	Research Design	Research Techniques
1	Creating a Flooding Intrusion Mitigation mechanism in SIP-based networks using a sorted Galloping Search algorithm	Constructive and Empirical	Experimental	Cross-sectional, before-and-after and Experiment	Quantitative and Scientific -laboratory experimentation
2	Preventing Fraudulent activity against SIP-based DoS flooding attacks using a sorted Quadratic Search algorithm	Constructive and Empirical	Experimental	Cross-sectional, before-and-after and Experiment	Quantitative and Scientific -laboratory experimentation
3	An extended Genetic Algorithm based Intrusion detection and prevention System against SIP-based Flooding attacks	Constructive and Empirical	Experimental	Cross-sectional, before-and-after and Experiment	Quantitative and Scientific -laboratory experimentation
4	A novel deep learning neighborhood technique to mitigate intentional DoS flooding attacks in VoIP systems	Constructive and Empirical	Experimental	Cross-sectional, before-and-after and Experiment	Quantitative and Scientific -laboratory experimentation
5	A deep analysis intruder tracing method to combat SIP-based Flooding attacks	Constructive and Empirical	Experimental	Cross-sectional, before-and-after and Experiment	Quantitative and Scientific -laboratory experimentation

3.2. Data Collection Techniques

The best way to efficiently prepare the collected data via the set-up testbed is to give broad information on the problem statement and the research questions. Since this work was constrained to authentication, the next step was to elaborate more on these threats. The general steps to solve the problems are as depicted in Fig. 2.

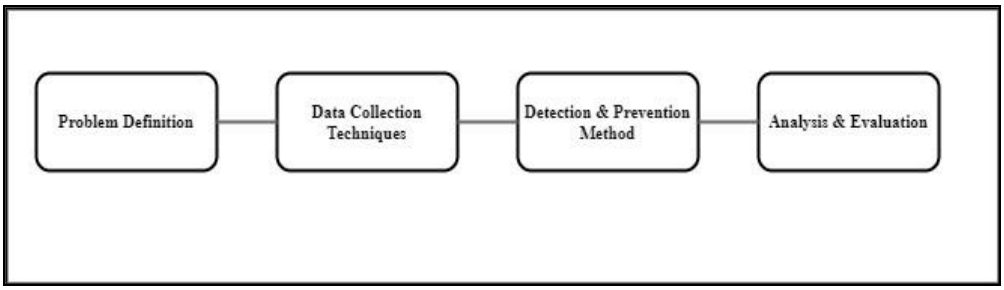

Fig.2. A basic walkthrough of the methodology.

The following are some questions that were addressed:

1) Who are the attackers and what are their goals?
2) What methods can be used to launch the attack?
3) What attack tools and/or written codes are used to launch the attacks?
4) What kinds of datasets were used in previous work done by researchers and what were the expected results obtained?
5) How did previous researchers prepare their datasets?

After obtaining these answers, the next phase was the method used to prepare the dataset. For regular IP networks, there exist several benchmark datasets [21] to evaluate the performance of intrusion and prevention approaches. However, except for Bad packets LLC [22], it has been observed that there are no public benchmark datasets available for VoIP networks due to privacy concerns. This renowned organization has logged over 8000 SIP attacks coming from 401 unique IP addresses. In general, researchers prepare their datasets using their set-up testbeds. In this study, the dataset from Bad Packets was taken into consideration when preparing the dataset.

3.3. Experimental Testbed

Fig.3 depicts the proposed three-layer defense system and testbed using a firewall, the proposed IDPS, and a setup Monitoring and Analysis Station (MAS) for an attempt to mitigate DoS attacks and DDOS attacks. The MiniSIPServer was used as it could be built-in for small and medium-size organizations of up to 500 users. The clients were the Zoiper, the LinPhone, and the MiniSIPphone softphone applications installed on various PCs and mobile phones.

The Firewall receives the SIP call message requests from the internal users and external users. It is set to filter only SIP call requests in a timeframe of 500 ms and consolidated to a CSV file with the tuples (Time, Source, Destination, Protocol & Info) to send the file to the IDPS for inspection at every 500 ms. The IDPS filters the detected

attacks and the corresponding source IP addresses are been quarantined and then sent to the Firewall to be added to its IP table for further filtering at its level when new bulks of messages enter the first layer of defense. The filtering of the quarantined IP addresses is been done before sending the legitimate packets to the SIP server. In the meantime, deep filtering is conducted by the third layer of defense. The IDPS jointly with the MAS conducts deep filtering on the legitimate dataset with broader time frames to identify false positive and false negative messages as well the users who are frequently been targeted in the eyes of the attackers. After deep filtering analysis, the IDPS may or may not detect more attacks and the corresponding IP addresses if detected are again sent to the Firewall for blocking.

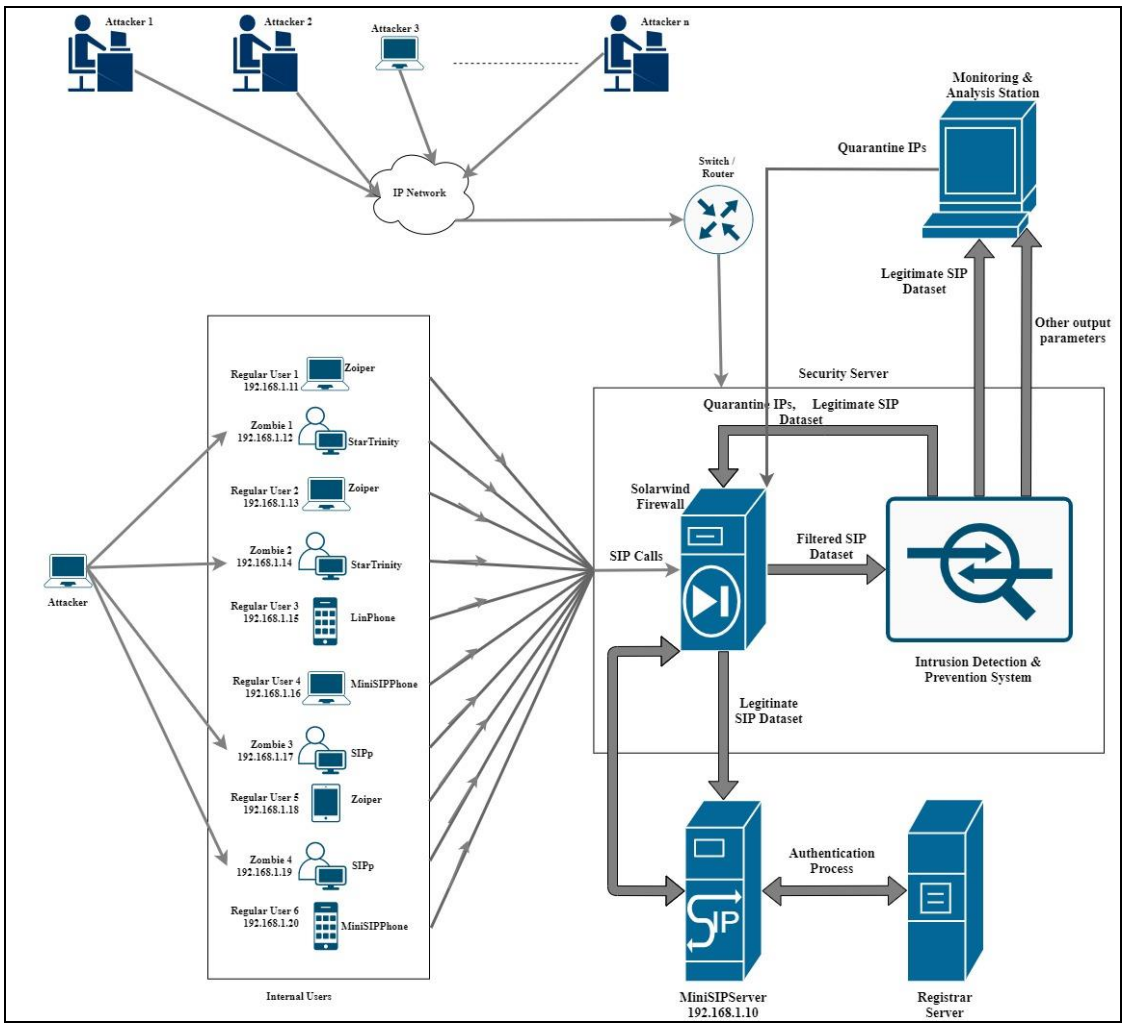

Fig.3. The proposed three-layer defense system.

4. Comparative Analysis

This section provides the results for the nine mitigation methods carried out in this study. A series of analyses are conducted to determine the best performing IDPS. The results depicted in this section are however a summary after pivoting other tabulated figures to reorganize the data

4.1. Comparison analysis within the SGP Group

Henceforth, Tables 2 & 3 depict the resultant information and contain the mean values and other statistical metrics (minimum, maximum, standard deviation) for the six experiments conducted at a different number of messages sent in 16 seconds. It is noted that all models use around 50 bytes of memory to process any number of messages varying between 348 and 8030 messages with h-SGP having the lowest mean value (49.99 bytes). Also, its mean value is been deviated by the smallest value (0.17 bytes) compared to other models. On the contrary, the m-SGP is better than the h-SGP in terms of processing time by 142 milliseconds. The deviation factor is also better for the m-SGP model. The q-SGP model is relatively the weakest model based on the poor results obtained. For the remaining KPIs, the performance is the same for all the models.

Therefore, overall, both the m-SGP and the h-SGP can be considered for future analysis.

Table 2. Statistical Metrics of SGP methods (Part 1)

Performance Parameters	Minimum			Maximum		
	m-SGP	h-SGP	q-SGP	m-SGP	h-SGP	q-SGP
Memory usage (B)	49.67	49.85	49.96	50.27	50.21	50.38
Processing Time (ms)	1100.81	875	1156.25	1265.62	2078.12	1562.5
Sensitivity (%)	98.00	98.00	98.00	98.00	98.00	98.00
Specificity (%)	98.00	98.00	98.00	100.00	100.00	100.00
Precision (%)	100.00	100.00	100.00	100.00	100.00	100.00
F-measure (%)	98.00	98.00	98.00	99.00	99.00	99.00
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.44	98.44	98.44	98.70	98.70	98.70
Detection Rate (%)	100.00	100.00	100.00	100.00	100.00	100.00

Table 3. Statistical Metrics of SGP methods (Part 2)

Performance Parameters	Mean			Standard Deviation		
	m-SGP	h-SGP	q-SGP	m-SGP	h-SGP	q-SGP
Memory usage (B)	50.05	49.99	50.16	0.21	0.17	0.20
Processing Time (ms)	1161.46	1304.20	1289.06	71.68	417.01	163.80
Sensitivity (%)	98.00	98.00	98.00	0.00	0.00	0.00
Specificity (%)	99.67	99.67	99.67	0.82	0.82	0.82
Precision (%)	100.00	100.00	100.00	0.00	0.00	0.00
F-measure (%)	98.83	98.83	98.83	0.41	0.41	0.41
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.66	98.66	98.66	0.11	0.11	0.11
Detection Rate (%)	100.00	100.00	100.00	0.00	0.00	0.00

4.2. Comparison analysis within the SQP Group

An in-depth analysis is further conducted to further verify the results as represented in Tables 4 and 5. All models require on average 50 bytes to process a flood of messages varying between 348 messages (small flood) and 8030 messages (very large flood). To detect fake messages in the floods, the q-SQP model is the fastest and requires 1215.49 milliseconds. However, a difference of 0.66 milliseconds in time is obtained when compared with the value obtained by the h-SQP, which is very tiny. For the other KPIs, the m-SQP shows deficiencies since the results are poorer than the two other models. For instance, its accuracy is 93.29% which is on the lower side while the false positive rate is quite high (6.5%). Similarly, the sensitivity and specificity are quite low for an IDPS as compared to the other two models (h-SQP and q-SQP) which have a good performance of 98% and 100% respectively. A zero false-positive rate confirms that the h-SQP and q-SQP are appropriate.

Table 4. Statistical Metrics of SQP methods (Part 1)

Performance Parameters	Minimum			Maximum		
	m-SQP	h-SQP	q-SQP	m-SQP	h-SQP	q-SQP
Memory usage (B)	50.02	49.82	49.83	50.28	50.28	50.33
Processing Time (ms)	1113.28	984.38	1062.50	1359.38	1359.38	1296.87
Sensitivity (%)	92.00	98.00	98.00	93.00	98.00	98.00
Specificity (%)	92.00	100.00	100.00	95.00	100.00	100.00
Precision (%)	94.00	100.00	100.00	96.00	100.00	100.00
F-measure (%)	93.00	98.00	98.00	94.00	99.00	99.00
False Alarm rate (%)	6.00	0.00	0.00	9.00	0.00	0.00
Accuracy (%)	91.98	98.44	98.44	93.75	98.73	98.73
Detection Rate (%)	94.38	100.00	100.00	95.94	100.00	100.00

Table 5. Statistical Metrics of SQP methods (Part 2)

Performance Parameters	Mean			Standard Deviation		
	m-SQP	h-SQP	q-SQP	m-SQP	h-SQP	q-SQP
Memory usage (B)	50.14	50.00	50.08	0.10	0.17	0.20
Processing Time (ms)	1258.47	1216.15	1215.49	102.42	147.38	94.86
Sensitivity (%)	92.50	98.00	98.00	0.55	0.00	0.00
Specificity (%)	94.00	100.00	100.00	1.10	0.00	0.00
Precision (%)	95.50	100.00	100.00	0.84	0.00	0.00
F-measure (%)	93.67	98.83	98.83	0.52	0.41	0.41
False Alarm rate (%)	6.50	0.00	0.00	1.22	0.00	0.00
Accuracy (%)	93.29	98.68	98.68	0.67	0.12	0.12
Detection Rate (%)	95.52	100.00	100.00	0.60	0.00	0.00

In the nutshell, after conducting a deep analysis, it is found that both h-SQP and q-SQP satisfy the requirement set for evaluation and hence can be considered for further analysis.

4.3. Comparison analysis between the Galloping and the Quadratic methods

A detailed study was conducted by computing the mean and the standard deviation as depicted in Tables 6 and 7. Again, the h-SGP model shows better performance since it requires less memory (49.99 bytes) to process messages followed by the q-SQP model (50 bytes), of which the difference is very minute by 0.01 bytes. Again, as mentioned in the preliminary analysis, the m-SGP shows excellent results (1160.03 milliseconds) compared to the other models. The quadratic models give excellent results for the other performance parameters. Therefore, considering the preliminary and the in-depth analyses, the q-SQP has the best score.

Table 6. Statistical Metrics of the Retaining Galloping and Quadratic methods (Part 1)

Performance Parameters	Minimum				Maximum			
	m-SGP	h- SGP	h- SQP	q- SQP	m-SGP	h- SGP	h- SQP	q- SQP
Memory usage (B)	49.67	49.85	49.83	49.82	50.27	50.21	50.33	50.28
Processing Time (ms)	1100.81	875.00	1062.50	984.38	1265.62	2078.12	1296.87	1359.38
Sensitivity (%)	98.00	98.00	98.00	98.00	98.00	98.00	98.00	98.00
Specificity (%)	98.00	98.00	100.00	100.00	100.00	100.00	100.00	100.00
Precision (%)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
F-measure (%)	98.00	98.00	98.00	98.00	99.00	99.00	99.00	99.00
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.44	98.44	98.44	98.44	98.70	98.70	98.73	98.73
Detection Rate (%)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 7. Statistical Metrics of the Retaining Galloping and Quadratic methods (Part 2)

Performance Parameters	Mean				Standard Deviation			
	m-SGP	h- SGP	h- SQP	q- SQP	m-SGP	h- SGP	h- SQP	q- SQP
Memory usage (B)	50.05	49.99	50.08	50.00	0.21	0.17	0.20	0.17
Processing Time (ms)	1160.03	1302.77	1215.49	1216.15	71.68	417.01	94.86	147.38
Sensitivity (%)	98.00	98.00	98.00	98.00	0.00	0.00	0.00	0.00
Specificity (%)	99.67	99.67	100.00	100.00	0.82	0.82	0.00	0.00
Precision (%)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00
F-measure (%)	98.83	98.83	98.83	98.83	0.41	0.41	0.41	0.41
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.66	98.66	98.68	98.68	0.11	0.11	0.12	0.12
Detection Rate (%)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00

4.4. Comparison analysis between Deep Analysis, Deep Learning and Genetic Algorithm Methods

An in-depth analysis is conducted by processing the measured values to obtain other statistical metrics (mean, standard deviation, etc.) which are tabulated in Tables 8 and 9. It is observed that the DAIT model uses a minimum of 49.76 bytes and a maximum of 50.42 bytes to process flood attacks, that is, it uses on average 50.03 bytes with a small standard deviation of 0.26 bytes. However, it can be noted that the e-GAP is also giving quite a good result in terms of memory utilization (50.10 bytes) with a relatively smaller deviation factor of 0.21 bytes. The CAN model is slightly inferior to the other two models and hence, such difference can be considered negligible. The DAIT model which shows a processing time of 738.83 ms has excellent value among the three models. Also, the mean accuracy for DAIT and e-GAP are very good (98.69%). Therefore, after conducting a deeper analysis, it can be observed that the figures obtained for e-GAP and CAN are relatively good and comparable to the DAIT model.

Table 8. Statistical Metrics of Deep Analysis, Deep Learning and Genetic Algorithm methods (Part 1)

Performance Parameters	Minimum			Maximum		
	DAIT	e-GAP	CAN	DAIT	e-GAP	CAN
Memory usage (B)	49.76	49.92	49.87	50.42	50.48	50.39
Processing Time (ms)	656.25	671.87	673.83	765.62	859.38	782.20
Sensitivity (%)	98.00	98.00	98.00	98.00	98.00	98.00
Specificity (%)	100.00	100.00	100.00	100.00	100.00	100.00
Precision (%)	100.00	100.00	100.00	100.00	100.00	100.00
F-measure (%)	98.00	98.00	98.00	99.00	99.00	99.00
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.44	98.44	98.44	98.75	98.75	98.70
Detection Rate (%)	100.00	100.00	100.00	100.00	100.00	100.00

Table 9. Statistical Metrics of Deep Analysis, Deep Learning and Genetic Algorithm methods (Part 2)

Performance Parameters	Mean			Standard Deviation		
	DAIT	e-GAP	CAN	DAIT	e-GAP	CAN
Memory usage (B)	50.03	50.10	50.13	0.26	0.21	0.22
Processing Time (ms)	738.93	757.82	742.51	42.25	81.34	45.49
Sensitivity (%)	98.00	98.00	98.00	0.00	0.00	0.00
Specificity (%)	100.00	100.00	100.00	0.00	0.00	0.00
Precision (%)	100.00	100.00	100.00	0.00	0.00	0.00
F-measure (%)	98.83	98.83	98.83	0.41	0.41	0.41
False Alarm rate (%)	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy (%)	98.69	98.69	98.66	0.12	0.12	0.11
Detection Rate (%)	100.00	100.00	100.00	0.00	0.00	0.00

To verify the validity and originality of the data collected at the output of each IDPS model, the experiments were repeated using cross-sectional and before-and-after. In this sequence, it was found that when the same data and the same number of attacks are injected at different times, the same output is obtained. And, secondly, when the same number of attacks and different messages are induced, almost the same outcome is obtained. The repetitive control of this experiment has determined that the threat to validity is almost negligible.

5. Discussion

In this study, the message traffic consists of a normal one mixed with the illegitimate one which is then captured by the application layer firewall and placed in a CSV file for analysis by the nine proposed flood mitigation models. During the test phase, it is observed that all models can separate legitimate messages from illegitimate ones. However, the results reveal that one model is better than the others when comparing the several performance indicators. The IDPS are complementing the work not only for the SIP server but also for the firewall. Indeed, the firewall has other jobs to address in a network. By simply redirecting the SIP messages to the prevention models, the firewall prevents its buffer from filling up as well as reduces the CPU utilization.

Concerning the results obtained following the different experiments conducted, all flood mitigation models use around 50 bytes on average to process messages during a silent flood behaviour or a very high flood behaviour. This can be observed in Fig.4 when comparing all models, which demonstrates that variations are almost the same. However, for the models designed using regular galloping and quadratic techniques (despite being enhanced), their processing time on average is very large (1241 ms) as compared to the other AI-based prevention models (CAN, DAIT and e-GAP) which is computed to have an overall mean value of 746 ms, that is, the latter are more efficient by 66.3%. This is also visible when comparing the models in a graph (Fig. 5). The CAN model which is been inspired by the KNN algorithm and RNN is inferior to the DAIT and the e-GAP by a very tiny difference (e.g. 3.58 ms on average for processing time, 0.01% in terms of accuracy and 0.03% in terms of accuracy and). Therefore, as the last observation, the study confirms that the CAN algorithm can be as well considered as an excellent model.

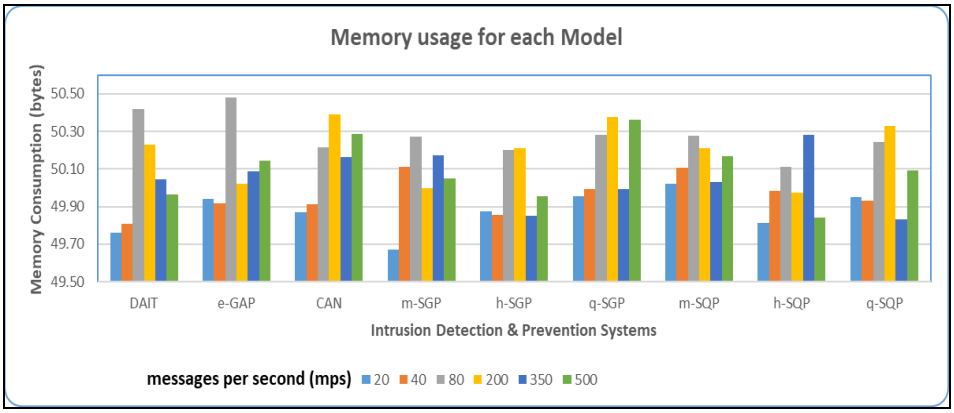

Fig.4. Analysis of Memory Usage for all models.

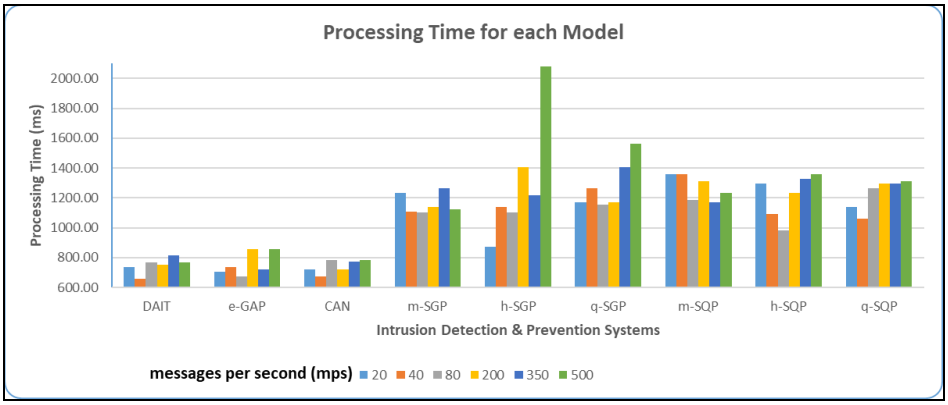

Fig.5. Analysis of Processing Time for all models.

An interesting result was obtained when analyzing the memory used by the DAIT and e-GAP for each slot of 500 ms for an attack duration of 16 seconds at two attack rates. This experiment is to investigate whether the system requires additional memory since more messages are entering the IDPS for examination.

As depicted in Fig.6 and Fig.7, for both IDPS and both attack rates, there is an increase (ramp up exponentially) in memory requirement until around 4.5 seconds (9 slots) and then the curves flatten afterward, that is, the system stabilizes after 4.5 seconds. The maximum memory which is required is around 50 bytes for most slots after the 9th slot (4.5 seconds). These results demonstrate that the memory will not get or hardly get exhausted irrespective of whether there is a small attack or high attack. The curve for the DAIT is slightly lower than that of the e-GAP, that is, there is a small deviation of 0.15 bytes and 15 bytes at 40 mps and 500 mps respectively.

Finally, the results of this study are compared with recent work conducted by Ahmed & Ali [7]. They argue that based on the existing literature, reports and publications indicate that most mitigation tools are inefficient since their detection rates are less than 95.5% and false-positive alarm of around 1.8%. Using a machine learning approach, Nazih et al. [5] have developed a signature and anomaly-based flood mitigation model and a malformed detector model both using SVM classifiers. They obtain a detection rate of 100% with a processing time/detection time of 0.73 ms per message tested. For the malformed mitigation model, the authors attain a detection rate of 100% and a processing time per message of 0.53 ms.

For this current study, the DAIT and the e-GAP models are having on average an F-score of 98.83%, a detection rate of 100%, a false rate of 0%, an accuracy of 98.7%, and finally a processing time per message of 0.092 ms and 0.094 ms respectively.

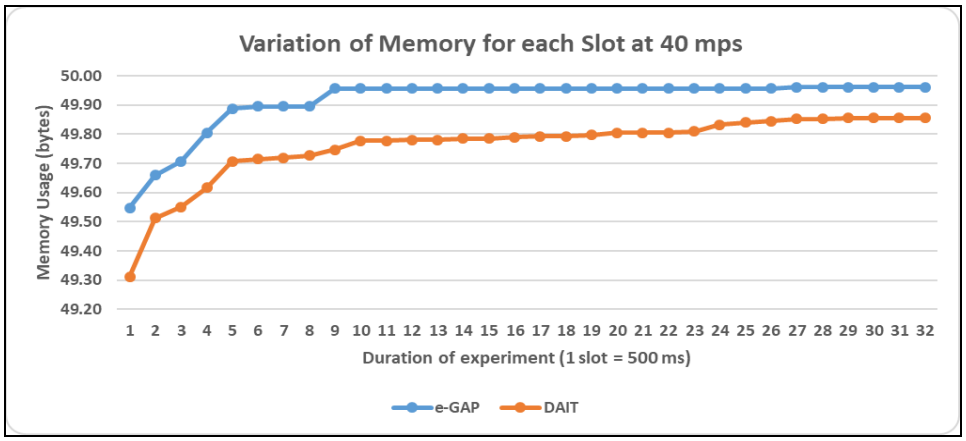

Fig.6. Amount of Memory used for each slot of 500 ms at 40 mps.

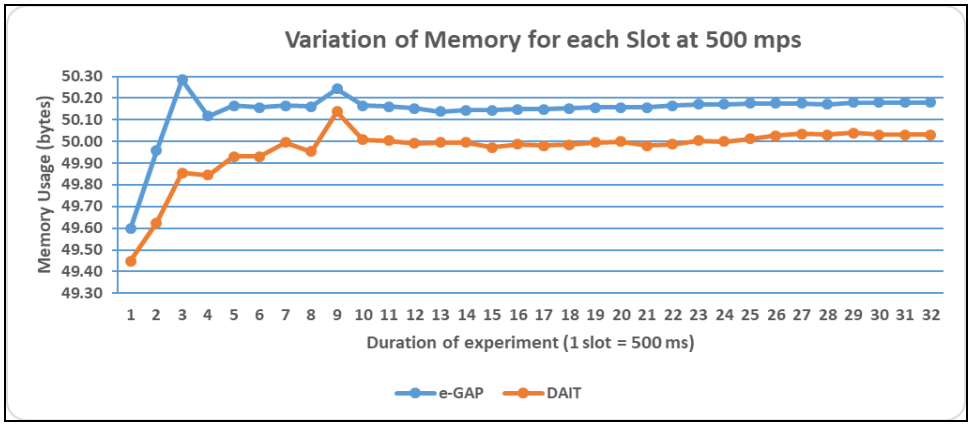

Fig.7. Amount of Memory used for each slot of 500 ms at 500 mps.

6. Conclusion

This section summarizes the major findings of this paper. In this study, nine models were evaluated and it was observed that all models detected genuine and non-genuine messages. However, it was noticed that each model gave slightly different values in relation to the nine key performance indicators. It is concluded that flood mitigation models use around 50 bytes on average to process messages during a flood attack irrespective of the size of the datasets. Based on the results, the DAIT and the e-GAP models performed the best among the nine flood mitigation tools while the regular algorithmic models showed the poorest results. Indeed, the processing time for the galloping and quadratic methods revealed to be very large (1241 ms) as compared to the other AI-based prevention models (CAN, DAIT and e-GAP). The latter had an overall mean value of 746 ms, which is, more efficient than the two formers by 66.3%. The CAN model was observed to give almost similar results as the DAIT and the e-GAP. Therefore, it is concluded that the CAN model could be used when building a network security system. The models were finally benchmarked with a very good machine learning signature and anomaly-based flood mitigation model using SVM classifiers to confirm their efficiency and efficacy. Results revealed that the three proposed flood mitigation models (CAN, DAIT & e-GAP) are more efficient than the SVM classifier machine learning models in terms of processing time.

Future work presumes an intense deep analysis of the nine IDPS algorithm mechanisms and also to execute the algorithms using different methodologies. Our current models do not cater for composite attacks. i.e, multiple attacks to the system at the same time. The models can be enhanced by strengthening the machine learning and related algorithms. Many other tests, experiments and deeper investigations have been left for future work due to the lack of availability to access and test on real industry-based datasets.

References

[1] Global Insights, “Insights to Innovation”, Available: https://www.gminsights.com/industry-analysis/voice-over-internet-protocol-voip-market. [Accessed 25 May 2021], 2021.

[2] A. Chauhan, N. Mahajan, H. Kumar S. Kaushal, “Analysis of DDoS Attacks in Heterogeneous VoIP Networks: A Survey”, *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, No. 6, pp 242-246, 2019.

[3] S. Ehlert, G. Zhang, D. Geneiatakis, G. Kambourakis, T. Dagiuklas, “Two Layer Denial of Service prevention on SIP VoIP

infrastructure", *Journal of Computers Communications, Elsevier*, Vol. 31, pp 2443–2456, 2008.
[4] M. A. Azad, R. Morla, K. Salah, "Systems and methods for SPIT detection in VoIP: Survey and future directions", *Journal of Computers & Security, Elsevier*, Vol. 77, pp. 1-20, 2018.
[5] N. Waleed, H. Yasser, E. Wail, A. Tamer, F. Hossam, "Efficient Detection of Attacks in SIP Based VoIP Networks Using Linear L1-SVM Classifier", *International Journal of Computers Communications & Control*, Vol. 14, Np. 4, pp 518–529, 2019.
[6] CiscoReport, "Cisco Annual Internet Report, 2018–2023 - Global Internet adoption and devices and Connections", [Accessed 27 May 2021], 2021.
[7] M. R. A. Ahmed, F. M. A. Ali, "Enhancing Hybrid Intrusion Detection and Prevention System for Flooding Attacks Using Decision Tree", *In: International Conference on Computer, Control, Electrical, and Electronics Engineering*, 2019.
[8] Y. Chen, "Detecting DoS attacks on SIP systems. *In the IEEE Workshop on VoIP Management and Security*", 2006.
[9] W. Li, W. Guo, X. Luo, X. Li, "On Sliding Window Based Change Point Detection for Hybrid SIP DoS Attack", *In IEEE Asia-Pacific Services Computing Conference*, 2010.
[10] X. Wan, Z. Li , Z. Fan, "A SIP DoS flooding attack defense mechanism based on priority class queue", *In the IEEE International Conference on Wireless Communications, Networking and Information Security*, 2010.
[11] B. Dhak, S. Lade, "An Evolutionary Approach to Intrusion Detection System using Genetic Algorithm", *International Journal of Emerging Technology and Advanced Engineering*, Vol. 2, No. 12, pp 632-36, 2012.
[12] W. Ahmad, D. Singh, "VoIP Security: A Model Proposed to Mitigate DDoS Attacks on SIP Based VoIP Network", *A Multi-Disciplinary Research Book*, pp 37-48, 2018.
[13] M. A. Ali, M. Farooq, "Application of Evolutionary Algorithms in Detection of SIP-based Flooding Attacks", *In the Annual Conference on Genetic and evolutionary computation, ACM*, pp 1419–1426, 2009.
[14] J. Tang, Y. Cheng, H. Yong, "Detection and prevention of SIP flooding attacks in voice over IP networks", *Proceedings of IEEE INFOCOM*, pp 1161-1169, 2012.
[15] S. Armoogum, N. Mohamudally, "Sorted Galloping Prevention Mechanisms against Denial of Service Attacks in SIP-based Systems", *In Proceedings of the 5th International Conference on Advanced Computing and Intelligent Engineering. Springer Nature, Elsevier Scopus*, 2020, ISSN: 2194-5357, DOI: 10.1007/978-981-33-4299-6,
[16] S. Armoogum, N. Mohamudally, "Prevention of fraudulent activities against SIP-based flooding attacks using extended sorted quadratic algorithms", *In Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications. Association for Computing Machinery (ACM), Elsevier Scopus*, 2020, Article 25, 1–7. DOI: https://doi.org/10.1145/3415088.3415113.
[17] S. Armoogum, N. Mohamudally, "An Extended Genetic Algorithm based Prevention System against DoS/DDoS Flood attacks in VoIP Systems", *In Proceedings of the 5th International Conference on Advanced Computing and Intelligent Engineering. Springer Nature, Elsevier Scopus*, 2020, (ISSN: 2194-5357), DOI: 10.1007/978-981-33-4299-6.
[18] Encyclopaedia-Britannica, "Survival of the fittest," [Online]. Available: https://www.britannica.com/science/survival-of-the-fittest. [Accessed 21 May 2021], 2021.
[19] S. Armoogum, N. Mohamudally, "Closest Adjacent Neighbour: a novel deep learning intruder detection technique in VoIP networks" *In Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications. Association for Computing Machinery (ACM), Elsevier Scopus*, 2021, Article 41, 1–7. DOI: https://doi.org/10.1145/3415088.3415129.
[20] S. Armoogum, N. Mohamudally, "A Novel Prevention Technique using Deep Analysis Intruder Tracing with a Bottom-up Approach against Flood Attacks in SIP-based Systems", Submitted in *Information and Computer Security*, 2021.
[21] I. Sharafaldin, A. Gharib, A. H. Lashkari, A. Ghornani, "Towards a Reliable Intrusion Detection Benchmark Dataset", *Journal of Software Networking*, pp 177-200, 2017.
[22] Bad Packets, "Meaningful Intelligence for an Evolving Cybersecurity Landscape", Retrieved May 27, 2021, from https://badpackets.net/, 2021.

Authors' Profiles

Sheeba. Armoogum is a Senior Lecturer at the Department of Information and Communication Technologies (ICT) at the University of Mauritius (UoM) and has more than 15 years of experience in teaching & learning at the tertiary level with several publications in top journals and conferences. She has completed her PhD in Cybersecurity from the University of Technology, Mauritius (UTM). Her fields of research & study are Cybersecurity, Cyber forensics, Networking and Security, Artificial Intelligence & Machine Learning, Software Engineering and Project Management. She is certified in Research Ethics and Research Ethics Evaluation by Training and Resources in Research Ethics Evaluation (TRREE). She was the past Head of the Department of Information and Communication Technologies at the UoM. She was part of several international conferences including the IEEE AFRICON 2013, IEEE EmergiTech 2016 and NEXTCOMP 2019 events.

Dr. Nawaz. Mohamudally is a Ph.D. holder from the University of Science and Technology Lille 1 (France) and is currently an Associate Professor at the University of Technology, Mauritius. Amongst others, he is the Lead Assessor of the Project Leadership Certification, a programme leading to Project Management Professional (PMP) CPDs. Dr. Mohamudally is an IBM Certified Mobile Development Instructor and recipient of several local and international research and development grants including from the MRIC. Moreover, he is a pioneer in innovative mobile app development and promotion in Mauritius.

How to cite this paper: Sheeba. Armoogum, Nawaz. Mohamudally, "A Comprehensive Review of Intrusion Detection and Prevention Systems against Single Flood Attacks in SIP-Based Systems", International Journal of Computer Network and Information Security(IJCNIS), Vol.13, No.6, pp.13-25, 2021. DOI: 10.5815/ijcnis.2021.06.02