# Data Stream Intrusion Alert Aggregation for Distributed Heterogeneous Sources

M.Thangavel<sup>1</sup> and Dr. P.Thangaraj<sup>2</sup>

Abstract— In the fast growing internet commercial transaction base, attacks on Internet infrastructure, anomaly intrusion traffic attacks combined with traditional network intruders, have become one of the most serious threats to the network security. The proposed system of the traffic anomaly detection method is carried out on the principle traces of non intrusive packet header data (statistical wavelet transform) obtained from the internet server traffic basement. Traffic is monitored at regular intervals to obtain a signal that can be analyzed through statistical techniques and compared to historical norms to detect anomalies. The proposed methodology of anomaly intrusion traffic detection envisions statistical non intrusive wavelet transform mechanisms for real-time data source extracted from NetCon server (Internet Service Provider popularly running at Erode Region) over a period of three months at various time intervals.

The experimental results suggest that address spoofing by attackers, which imply that such attacks will be invisible to indirect backscatter measurement techniques. Further, at the detailed packet-level characterization (e.g., attack destination ports), there are significant differences between anomaly and traditional intrusion attack measurements. Thus, there is tremendous value in moving towards direct observations to better understand recent intrusion attacks. Traffic Anomaly intrusion measurements additionally provide information inaccessible to traditional network intrusion measurements, enabling us to better understand how to defend against attacks.

*Index Terms*—Traffic Anomaly, Intrusion Detection, Server Intruders, Traffic Attack and NetCon

#### I. INTRODUCTION

Many approaches have been studied to detect, prevent and mitigate malicious network traffic. For example, rule-based approaches, such as IDS (intrusion detection system), try to apply previously established rules against incoming traffic to detect and identify potential Intrusion attacks close to the victim's network. To cope with novel attacks, however, IDS tools such as Snort require to be updated with the latest rules. This project looks at the problem of designing generalized measurement based real-time detection mechanisms.

Measurement-based studies have considered traffic volume [3], number of flows as potential signals that can be analyzed in order to detect anomalies in network traffic, while the system further treat the traffic headers such as addresses and port numbers. Input data from multiple sources (i.e., all links in a network), while the work focuses on a single link at a time. Some approaches proactively seek

methods to suppress the overflow of traffic at the source [5]. Controls based on rate limits have been adopted for reducing the monopolistic consumption of available bandwidth, to diminish the effects of attacks, either at the source or at the destination [5], [7], [12]. The apparent symptoms of bandwidth attack may be sensed through monitoring bit rates [10] and/or packet counts of the traffic flow. Bandwidth accounting mechanisms have been suggested to identify and contain attacks [8, 9, and 11]. Packeteer [2] and others offer commercial products that can account traffic volume along multiple dimensions and allow policy-based rate control of bandwidth.

Pushback mechanisms have been proposed to contain the detected attacks closer to the source [11, 12]. Trace back has been proposed to trace the source of Intrusion attacks even when the source addresses may be spoofed by the attacker. However, sophisticated low-rate attacks do not give rise to noticeable variance in traffic volume, could go undetected when only traffic volume is considered. Recently statistical analysis of aggregate traffic data has been studied. In general, the generated signal can be analyzed by employing techniques such as FFT (Fast Fourier Transform) and wavelet transforms. FFT of traffic arrivals may reveal inherent flow level information through frequency analysis. Fourier transforms and wavelets have been applied to network traffic to study its periodicity.

The previous work in [1] and the work in [3] studied traffic volume as a signal for wavelet analysis and these earlier studies have considerably motivated the current study. The study builds on this earlier work and extends the statistical analysis of traffic data further in analyzing other packet header data, such as addresses and port numbers in real-time. Various forms of signatures have been traditionally utilized for representing the contents or identities of documents.

This earlier work motivated the representation of aggregate network traffic data in a compact data structure. A similar data structure was employed in [4], with significant differences in processing of collected data, detection mechanisms and the resulting traffic anomaly detectors. The structure of addresses at various points in the network was observed to be multi-fractal in [6].

## II. RELATED WORKS

Tools for analyzing network traffic allow the detection of anomalies in the environment, including attacks and unusual events in the network, and enable fast execution of actions to avoid that the detected threats can propagate through the network. The network traffic should be monitored in regular intervals to obtain data that will be analyzed by statistical or intelligent techniques in search of anomalies. The idea is storing data from normal traffic (historical data) for future comparison with real traffic (current data) to detect eventual anomalies. By observing the traffic and correlating it to its previous states, it may be possible to see whether the current traffic is behaving in a similar / correlated manner [12]. This approach is named anomaly detection.

Anomalies on the network traffic are defined as previously unseen (yet legitimate) traffic behaviors. A wide range of unusual events some of which, but not all, may be malicious known as traffic anomalies are commonplace in today's computer networks. Identifying, diagnosing and treating anomalies such as failures and attacks in a timely fashion are a fundamental part of day to day network operations. Operators need to detect these anomalies as they occur and then classify them in order to choose the appropriate response. The principal challenge in automatically detecting and classifying anomalies is that anomalies can span a vast range of events: from network abuse (e.g., Intrusion attacks, scans, worms) to equipment failures (e.g., outages) to unusual customer behavior (e.g., sudden changes in demand, flash crowds, high volume flows), and even to new, previously unknown events. A general anomaly diagnosis system should therefore be able to detect a range of anomalies with diverse structure, distinguish between different types of anomalies and group similar anomalies.

Regardless of whether the anomalies in question are malicious or unintentional, it is important to analyze them for two reasons i.e., one, anomalies can create congestion in the network and stress resource utilization in a router, which makes them crucial to detect from an operational standpoint, and second one is that some anomalies may not necessarily impact the network, but they can have a dramatic impact on a customer or the end user. Existing postmortem analysis [4] of traffic volume (in bytes) can reveal changes in traffic patterns. Flow Scan and Auto Focus [5] are used as off-line traffic analyzing tools. Rigorous real-time analysis is required for detecting and identifying the anomalies so that mitigation action can be taken as promptly as possible. Analyze the effectiveness in real-time analysis of traffic data. Real-time analysis provides means of online detection of anomalies while they are in progress. Real-time analysis employs smaller amounts of data in order to keep such analysis simple and efficient. Real-time analysis also requires that any indications of attacks or anomalies be provided with short latencies. This trade-off between robustness and latency makes real-time analysis more challenging.

The proposed approach look at aggregate packet header data to improve scalability and to effectively deal with anomalies in cases where individual attack flows may not look anomalous. Use discrete wavelet transform model to correlate destination IP address. With discrete wavelet outcome, use statistical analysis for effective detection of anomalies. Provide a multidimensional indicator using the correlation of port numbers and the number of flows as a means of detecting anomalies. Results from trace-driven evaluation suggest that proposed approach could provide an effective means of detecting anomalies close to the source.

A wireless IDS can be used in one of two ways centralized or decentralized. In a decentralized environment each WIDS operate independently, logging and alerting on its own. Beside this also means that wherever WIDS must be managed. In a large network this can quickly become overwhelming and inefficient and therefore not recommended for networks with more than one or two access points. The idea behind a centralized WIDS is that the sensors are deployed that relate information back to one central point. This point could send alerts and event logging and serve as a single point of administration for all sensors. Another advantage of a centralized approach is that the sensors can work together to capture a wider range of events more closely (Seamans and Yang, 2004). In this approach, there are three possibilities, the sensors can be used. The first is by using existing Access Points (AP).

Some accesses points are on the market are able to simultaneously function as an access point and WIDS sensor. This option has the potential to be cheaper than others, however there is a downside. Using the AP for both functions will reduce the performance, potentially creating a "bottle neck" on the network. The second option is the use of "dumb" sensors. These devices simply relay all information to the central server and rely on the server for all events. Although inexpensive, all information in one place, the impact on the performance of the wired network and the creation of a single point of failure is sent to the server. The third option is the use of intelligent sensors. These devices actively monitor and analyze wireless traffic, identify attack patterns and rouge devices as well as look for deviations from the norm. Then these events are reported to a central server and allow administrators to invoke countermeasures.

There are three types of MAC 802.11 frames, data, control and management (Gupta and Shroff, 2010). The wireless IDS is unique in that it detects attacks against the 802.11 frame at layer two of the wireless network. Most wireless attacks target management frames, since they are responsible for authentication, association, dissociation, beacon and probe request/response (Tague *et al.*, 2011). Wireless threats such as attacks man-in-the-middle, rouge access points, war drivers and denial of service attacks function within the 802.11 frames and cannot be detected on layer three past access point. Wired IDS will not receive these frames, since the management frames are not forwarded to upper layers of the OSI model.

As traditional wired intrusion detection systems that are used to monitor a network, wireless intrusion detection systems need a dedicated interface. This wireless interface must be in monitor mode, the operator also known as RFMON mode, this mode is similar to promiscuous mode for installation materials and the device can accept incoming traffic (Gupta and Shroff, 2010). Another important aspect of wireless IDS is that the monitoring interfaces between the 12 channels available to wireless networks. Several wireless attacks work by utilizing a rogue access point on a different channel. For instance, attacks man-in-the-middle utilizes a rogue AP that at least 5 channels away from the target AP. Without channel hopping the wireless IDS would be blind to attacks that function on other channels.

## III. SYSTEM MODEL

## A. Traffic Sources

The proposed system model's traffic source workable data are obtained from the routers deployed in the network domain considered for evaluation. A traffic monitoring at a source network enables a detector to detect attacks early and is able to control hijacking of admin domain machines. Outbound filtering has been advocated for limiting the possibility of address spoofing i.e., to make sure that source addresses correspond to the designated addresses for the campus. With such filtering in place, we can focus on destination addresses and port numbers of the outgoing traffic for analysis purposes.

### B. Anomaly Detection mode

The anomaly detection mode is based on the functions of network traffic splitter, and statistical data transformation. The traffic splitter generates network traffic signal from packet header traces or data flow records. The statistical data transformation analysis is carried out with wavelet transforms of IP address and port number correlation over several timescales. Then the detection of attacks and anomalies are checked using thresholds. The analyzed information will be compared with historical thresholds to see whether the traffic's characteristics are out of regular norms. This comparison will lead to some form of a detection signal that could be used to alert the network administrator of the potential anomalies in the network traffic.

## C. Data Transform

The generated signal can be, in general, analyzed by employing techniques such as FFT (Fast Fourier Transform) and wavelet transforms. The analysis carried out on the signal may exploit the statistical properties of the signal such as correlation over several timescales and its distribution properties. Since wavelet analysis can reveal scaling properties of the temporal and frequency dynamics simultaneously unlike Fourier Transform, proposal compute a wavelet transform of the generated address correlation signal over several sampling points. Through signal can be detected in certain timescales that imply frequency components, and in certain positions of the timescales that mean temporal information, can induce the frequency and temporal components respectively. Discrete Wavelet Transform (DWT) consists of decomposition and reconstruction. The proposed model iterates a multilevel one-dimensional wavelet analysis up to 8 levels in case of the postmortem analysis, so our final analysis coefficients. The filtered signal is down sampled by 2 at each level of the analysis procedure. The signal of each level has an effect that sampling interval extends 2 times. Consequently it means that the wavelet transform identifies the changes in the signal over several timescales. When use t minutes as sampling interval j, the time range at a level spans  $t^*2j$ minutes. This time range can independently sample and

#### restore frequency components

## D. Statistical Method Interaction

## Individual fields in the packet header are analyzed to

observe anomalies in the traffic. Individual fields in the traffic header data take discrete values and show discontinuities in the sample space. For example, IP address space can span 232 possible addresses and addresses in a sample are likely to exhibit many discontinuities over this space making it harder to analyze the data over the address space. In order to overcome such discontinuities over a discrete space, we convert packet header data into a continuous signal through correlation of samples over successive samples. To investigate the sequence of a random process, we employ a simplified correlation of time series for efficiency computational without compromising performance.

For each address in the traffic count the number of packets sent in the sampling instant. For computing address correlation signal, consider two adjacent sampling instants. The detection model define address correlation signal in sampling point. If an address spans the two sampling points i.e., n-1 and n, the user obtain a positive contribution. In order to minimize storage and processing complexity, employ a linked data structure. A location count is used to record the packet count for the address j in *ith* field of the IP address through scaling. This provides a concise description of the address instead of 232 locations that would be required to store the address occurrence uniquely.

The detection model filters this signal by computing a correlation of the address in two success samples. Consequently four correlation signals are calculated. The employment of this approximate representation of addresses allows us to reduce the computational and storage demands by an appreciable factor. In order to generate the address correlation signal at the end of sampling point, multiply each segment correlation with scaling factors. From a statistical view, they have an approximately same mean and dispersion standard deviation as cross-correlation coefficient.

The hybrid model for intrusion detection attack in wireless network selects the optimal features to specially detect 802.11 specific intrusion attacks. The hybrid model of feature selection used information gain ratio. It evaluates the relevance of 802.11 specific features and its measures of information gain ratio indicate the features essentially needed for detecting anomaly intrusion attack in the network. The k-means clustering is further applied to select the optimal feature set to improve the efficiency of the intrusion detection by classifying the attributes to the associated attack class obtained from the records sets of the tracked data from the wireless network MAC frames. Statistical characteristic of data traffic learning is further used to reduce the detection time and improve the overall performance of wireless intrusion detection with optimal set of features for detection criteria.

The improved version of the hybrid model is developed using existing wireless intrusion detection specific to 802.11 features, by utilizing integrated statistical and clustering principles. The clustering effectively identify the set of reduced optimal features. The optimal feature obtained is more efficient in detecting the intrusions of low level network framework. In addition simulations are carried out to compare the performance of existing classical intrusion detection schemes in wired network to the proposed hybrid network.

**De-authentication attack:** De-authentication attack is an easy to mount attack that can work on any type of 802.11 networks (WEP and WPA) which modifies the 802.11 MAC frame. It enables an attacker to terminate the connection of all stations connected to the wireless network. The attacker sends a de- authentication frame with a destination address. The stations that receive this frame will automatically disconnect from the network. The operation is repeated continuously to prevent the stations from maintaining their connections to the access point.

**802.11 MAC frame:** The 802.11 MAC frame consists of a MAC header, body frame and a Frame Check Sequence (FCS), as shown in Fig. 1.

The numbers on the Fig. 1 represents the number of bytes for each field. The frame control field contains control information used to define the nature of the 802.11 MAC frame and providing information necessary for the following fields to understand how to process the MAC frame. The numbers on the Fig. 2 represents the number of bits for each field of the control settings in the MAC Frame.

Each sub-field under frame control field is explained in Fig. 2 Protocol Version provides the current version of the 802.11 protocol used. Receiving STAs use this value to determine if the version of the protocol of the received frame is supported. Type and Subtype determines the function of the frame. There are three different frame type fields: control, data and management. There are multiple subtype fields for each frame type . Each subtype determines the specific function to perform for its associated frame type. To DS and From DS indicates whether the frame is going to or exiting from the Distributed System (DS) and is only used in data type frames of STAs associated with an AP.

Other fragments indicate whether other fragments of the frame, either data or management type, followed. Retry indicates whether or not the frame, for either data or management frame types, is being retransmitted. Power management specifies if the sending STA is in active mode or power-save mode. More data indicates to a STA in power-save mode, the AP has to send more frames. It is also used for access points, to indicate that additional broadcast/multicast frames to follow. WEP indicates whether or not encryption and authentication are used in the frame. It can be set for all data frames and management frames, which have the subtype set to authentication. Order indicates that all received data frames must be processed in order.

Transmitter Address (TA) shows the MAC address of the STA that transmitted the frame into the wireless medium. The frame body contains the data or information included in either management type or data type frames. STA transfer with a Cyclic Redundancy Check (CRC) in all areas of the MAC header and the area of the body structure to generate the FCS value. The receiving STA then uses the same CRC calculation of its own FCS field of the mixture distribution. The approach adopted is Maximum Likelihood (ML) estimation. Hybrid network anomaly intrusion detection is aware of the situation and tries to keep the model updated to current attack. There is a trade-off between the running time (reaction time) and accuracy. It is possible to decide upon the existence of a new attack instance when only one observation is made.

value to determine whether or not errors have occurred during the transmission.

**Feature extraction:** Feature extraction is responsible for the extraction of attributes and characteristics that are effective in detecting the interference of the

802.11 MAC frame fields. The attributes are selected depending on the type of the frame and the detection algorithm. This set of characteristics is sent to the local misuse detection and the central module for detecting anomalies.

Statistical anomaly traffic characteristic in MAC frames: MAC frame fields in the 802.11 packet header are analyzed to observe anomalies in the traffic. Individual fields in the traffic header data take discrete values and show discontinuities in the sample space. MAC address space span multiple addresses in a sample are likely to exhibit many discontinuities over this space. It is difficult to analyze the data over the address space. In order to overcome such discontinuities in a discrete space, the packet header data transform in to continuous signal through correlation of samples successive samples. To investigate the sequence of a random process, we use a simplified correlation time series for computational efficiency without sacrificing performance.

For each MAC address in the traffic, count the number of packets sent in the sampling instant. For computing address correlation signal, consider two statistical model anomaly detection filters this signal by calculating a correlation of the address in two samples of success. Consequently four correlation signals are calculated. This approximate representation of addresses allows us to reduce the demands of the computing and storage for a considerable factor. To create the correlation of signaling messages at the end of the sampling point, multiply each segment of the correlation with scaling factors. From a statistical point of view, they have more or less same mean and dispersion standard deviation as cross-correlation coefficient.

Cluster intrusion alert generation: Attributes in an event of data traffic in the hybrid network are independent of a particular attack instance used for clustering. Attributes are dependent on the instance of the attack used in the clustered alert aggregation process to distinguish different attach instances. Dependent metrics such as source IP address identify the attacker. Independent metrics, i.e., destination port -80, in case of web-based attacks. Both contain the attacker's real target services specifically for a particular service objective. Regarding, an attack instance, a random process generate alerts, that are distributed according to a certain multivariate probability distribution. The alert space is composed of several attributes. Reconstructing an attack situation from observed samples estimate all the parameters of the mixture distribution. The approach adopted is Maximum Likelihood (ML) estimation. Hybrid network anomaly intrusion detection is aware of the situation and tries to keep the model updated to current attack. There is a trade-off between the running time (reaction time) and accuracy. It is possible to decide upon the existence of a new attack instance when only one observation is made.

mixing coefficients at certain points in time.

## IV. EXPERIMENTAL EVALUATION

The experimental evaluation of anomaly traffic detection was conducted based on the monitored traces of network traffic generated synthetically as shown in Figure.1 and real time samples obtained from NetCon server of an Internet Service Provider at our Erode Region. The real trace samples of the NetCon server was carried for a period of one month connecting with 10Mbps broad band link. Additionally employ the packet traces from the monitored trace data and super-imposed with simulated virtual attacks. The scenario in which the samples taken from the NetCon server having nearly 23000 connections at the rate of 1MBPs 600 packets / second. These traces were anonym, but preserved IP prefix relationships. The deployment of the anomaly traffic detection applied in the network is shown in Fig 1. In this mode of the experiment, packets with different sizes flow from ISP to network.

Based on bandwidth threshold the admission of anomalous packets will be controlled by dropping the packets. Bandwidth threshold is applied on the packet examination is shown in Figure 2. This made the control scheme of the bandwidth mode to drop some unwanted packets. The dropped packet size is 1000 bytes. As shown in Figure-3 packets with less than 500 MB are allowed to transmit into the destination.

The attacks cover a diversity of behaviors and allow us to deterministically test the efficacy of proposed mechanisms. These are classified by following criteria of persistency, IP address, protocol, port and size. Persistency is the process of remnant attacks persistently assault. IP address attack damage a portion of addresses preserve the class-A and a partition of it. Comparison of full-32 bit correlation and data structure addresses preserve class-B for the infiltration efficiency. The other type is randomly generated address. The Protocol ICMP, TCP and UDP, are exploited in turn. The port attacks are a representative #80 that stands for the reserved ports for well-known services. The port targets for randomly generated destination ports that are used to probe port scan. The three size denominations are random ones The anomaly detection mode is based on the functions of network traffic splitter and statistical data transformation. The traffic splitter generates network traffic signal from packet header traces or data flow records. The statistical data transformation analysis is carried out with wavelet transforms of IP address and port number correlation over several timescales. Then the detection of attacks and anomalies are checked using thresholds. The analyzed information will be compared with historical thresholds to see whether the traffic's characteristics are out of regular norms. This comparison will lead to some form of a detection signal that could be used to alert the network administrator of the potential anomalies in the network traffic.





Figure.2 Packets with greater than 500MB are dropped



Figure.3 Packets with less than 500MB are sent

#### V. PERFORMANCE EVALUATION

Detection of anomalies using the real attack traces are based on composite approach with network traffic measures on the sampled time band. The weighted correlation signal of IP addresses is used for wavelet transform with real attacks. The anomaly detection results of0020the wavelet-transform and reconstructed signal are supposed to be evaluated from the experimentation.

If statistical parameters of network traffic, such as mean and standard deviation, are stationary distributed under given traffic, parameters of specific day could be applied to other days. The experiment is conducted over a given time scale of two months trace and analyze their statistical summary measures. Wide-sense stationary (WSS) regarding these traces are obtained with the average independent on time, and autocorrelation function of time difference regardless of sample path. Based on communication, the analysis of the ambient trace could be considered as WSS Gaussian white noise, on the other hand, the attack and anomaly could be considered as signal of interest. It illustrates the thresholds status over the sampled time scale on anomaly traffic range.



The graph1 depicts the output of the simulation by varying the number of nodes in the network traffic. There is an appreciable change in the throughput of the data communication. As the number of nodes increases, throughput decreases. By comparing it with non bandwidth threshold model, the throughput is high in the anonymous traffic detection with bandwidth threshold model.

Graph 2: Number of Nodes Vs Jitter



The simulation result based on the node variation affecting the jitter is depicted in the graph 2. The jitter value increases as the number of nodes increased in the anonymous traffic network. As for comparison made on existing method (without bandwidth threshold), the jitter is low in the proposed method.

#### DISCUSSION

The simulation of anomaly intrusion traffic detection was conducted based on the monitored traces of hybrid network (wired and wireless) traffic generated from real time data traffic from ISP servers. The real trace of samples was carried for a period of one month connecting with 10Mbps broad band link comprising of wireless and wired network servers. The samples taken from the ISP server have 1000s of wired and 1000s of wireless connection, at the traffic rate varying from 256Kbps to 1MBPs. These traces were anonym, but preserved with MAC and IP prefix relationships. The deployment of the anomaly traffic detection applied in the hybrid network is done based on the clustering Meta alert and statistical characteristics of data traffic. The simulation was conducted on IBM PC Compatible Machines with 2.20 GHz and 2GB of RAM in NS2 simulator.

The simulation evaluate the performance, such as the detection accuracy and communication overhead of our scheme through simulations. We use a field size of 1000×1000 m where 80 nodes are uniformly distributed. One stationary sink and one stationary source sit on opposite sides of the field, with about 2 to 3 hops in between. Carry out a simulation event in which the source generates 500 reports in total and one report is sent out every two seconds. Packets can be delivered hop-by-hop at 10 Kbps. In order to avoid detection, the malicious nodes drop only part of the packets passing by. To make our scheme more resilient in poor radio conditions, implement a hop-by hop transport layer retransmission mechanism. The retransmission limit is 5 by default. The channel error rate is 10% by default, which is usually regarded as a rather harsh radio condition. Each simulation runs 10 times and the result shown is an average of these runs.

The proposed metrics evaluate the detection accuracy and communication overhead of the existing and proposed schemes. Alarm reliability measures the ratio of the number of detected maliciously-dropped packets to the total number of lost packets detected including those lost due to poor radio conditions. Undetected rate measures the ratio of the number of undetected maliciously-dropped packets to the total number of maliciously-dropped packets.

Relative communication overhead measures the ratio of the total communication overhead in a system that incorporates our detection scheme against a system that does not.

The simulation conducted also performed evaluation of cluster alert aggregation approach. The simulation deployed different hybrid network data sets to demonstrate the feasibility of integrated statistical and cluster based anomaly intrusion model. Several weeks of training and test data have been generated on a test bed that emulates a small confidential data site. In the hybrid network scenario of both wired and wireless, the generated network traffic is analyzed for its intrusion characteristic detection. The simulation used both MAC frame and IP address dump as input data and analyzed varied attack instances against various target hosts. The extracted statistical information from the network traffic data, apply Support Vector Machines (SVM) to classify the clustered data characteristic samples. The performance graph shown in Fig. 3 shows that the Hybrid network anomaly intrusion detection proposed in this work has better response time for the detection rate compared to that of the classical intrusion detection scheme. The percentage of improvement is made nearly 17% in terms of response time for the anomaly intrusion detection from the hybrid network data traffic.



Graph 3: Performance of anomaly intrusion detection in hybrid network

The tabulated values of the detection rate response time against number of data records extracted from the hybrid network data traffic shown in Table 1 indicates that our proposal has better response time in detection anomaly intrusion both in wired and wireless network. With varying threshold to the data traffic rate of the clustered alert aggregation, the True Positive Rate (TPR, number of true positives divided by the sum of true positives and false negatives) and False Positive Rate (FPR, number of false positives divided by the sum of false positives and true negatives) are identified for the trained data traffic records.

Various operating points are marked to evaluate the normal traffic scenario and anomaly traffic intrusion at different data rates. The simulation conducted investigated, aggregation under idealized conditions where we assume to have a perfect detector layer with no missing and no false alerts at all. As attributes for the alerts, use MAC frame and IP address, source and destination port, the attack type and the creation time differences (based on the creation time stamps). The performance of the hybrid network intrusion detection shows improved detection rate with reduced set of features. The false positives rate identified is the percentage of frames containing normal traffic classified as intrusive frames, which is minimal in our scheme. False negatives rate identified is the

percentage of frames generated from wireless attacks which are classified as normal traffic which is precisely evaluated in our scheme.

# VI. CONCLUSION

The anomaly intrusion traffic detection work carried out in this paper combine multiple independent data sources to study combined traditional intrusion attack and anomaly intrusion and provides the statistical wavelet based detection mechanism. The examined backscatter data from a mostly unused NetCon server network along with flow anomaly based traffic data from a tier-1 ISP network. The attack characterization indicates that most properties such as attack duration, packet count, packet rate, and dominant protocol type match fairly well in the two data sets. In addition load, demand capacity of the server at lean and heavy traffic scenarios are observed to provide better clarity of anomaly intrusion detection across all time zone of the server uptime. Using anomaly intrusion attack measurements, the proposed model performed an analysis of several traffic anomaly properties which is impossible using traditional intrusion measurements.

There is significant predictability in attacks both in terms of their originating server as well as from which interface they enter a large ISP network. Small businesses seem to be the most common targets of attacks. These results have significant implications for attack defense. With respect to anomaly intrusion attack detection and understanding attacks, relying on traditional measurements is clearly not sufficient given current trends in attacks, since very few if any of attacks appear to be using spoofed source addresses. As an alternate, it is found that anomaly traffic intrusive attack measurements can provide significantly more diagnostic capability that can better guide the design and deployment of attack defenses. In fact, there are positive implications for attack defense. From the perspective of service providers, intrusion attacks are really not as distributed as they are made out to be. Since the vast majority of malicious traffic arises from a small set of internet attacking servers and network

The proposed metrics evaluate the detection accuracy and communication overhead of the existing and proposed schemes. Alarm reliability measures the ratio of the number of detected maliciously-dropped packets to the total number of lost packets detected including those lost due to poor radio conditions. Undetected rate measures the ratio of the number of undetected maliciously-dropped packets to the total number of maliciously-dropped packets.

Service providers can ensure significant protection for their customers with even limited (but intelligently targeted) deployment of intrusion defense mechanisms.

#### REFERENCE

- [1] A. Ramanathan, "WADeS: A tool for distributed denial of service attack detection" M.S. thesis, TAMU-ECE-2002-02, Aug. 2002.
- [2] NLANR measurement and operations analysis team, NLANR Network Traffic Packet Header Traces, Aug.

2002 [Online]. Available: http:// www.pma.nlanr.net/Traces/

- [3] P. Barford et al., "A signal analysis of network traffic anomalies," in ACM SIGCOMM Internet Measurement Workshop, Nov. 2002.
- [4] T. M. Gil and M. Poletto, "MULTOPS: A data-structure for bandwidth attack detection," in USENIX Security Symp., Aug. 2001.
- [5] J. Mirkovic, G. Prier, and P. Reiher, "Attacking DDoS at the source," in IEEE Int. Conf. Network Protocols, Nov. 2002.[6] E. Kohler, J. Li, V. Paxson, and S. Shenker, "Observed structure of addresses in IP traffic," in Proc. ACM IMW, Nov. 2002.
- [6] A. Garg and A. L. N. Reddy, "Mitigation of DoS attacks through QoS regulation," in Proc. IWQOS, May 2002.
- [7] Smitha, I. Kim, and A. L. N. Reddy, "Identifying long term high rate flows at a router," in Proc. High Performance Computing, Dec. 2001.
- [8] I. Kim, "Analyzing network traces to identify long-term high rate flows," M.S. thesis, TAMU-ECE-2001-02, May 2001.
- [9] Y. Zhang, L. Breslau, V. Paxson, and S. Shenker, "On the characteristics and origins of internet flow rates," in ACMSIGCOMM, Aug. 2002.
- [10] R. Mahajan et al., "Controlling high bandwidth aggregates in the network," ACM Comput. Commun. Rev., vol. 32, no. 3, Jul. 2002.
- [11] J. Ioannidis and S. M. Bellovin, "Implementing pushback: Router-based defense against DDoS attacks," in Proc. Networkand Distributed System Security Symp., Feb. 2002.
- [12] J. Mirkovic, S. Dietrich, D. Dittrich, and P. Reiher, Internet Denial of Service: Attack and Defense Mechanisms. Prentice Hall, 2005.
- [13] R. Beverly, iThe Spoofer Project: Inferring the Extent of Internet Source Address Filtering on the Internet Proceedings of USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI 2005), July 2005.
- [14] V. Sekar, N. Duf\_eld, J. van der Merwe, O. Spatscheck, and H. Zhang, i LADS: Large-scale Automated DDoS Detection System, USENIX Annual Technical Conference, 2006.
- [15] A. Kumar, V. Paxson, and N. Weaver, Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event, ACM IMC, 2005.
- [16] Gupta, G.R. and N.B. Shroff, 2010. Delay analysis for wireless networks with single hop traffic and general interference constraints. IEEE/ACM Trans. Network., 18: 393-405.
- [17] DOI: 10.1109/TNET.2009.2032181
- [18] Khoshgoftaar, T.M., S.V. Nath, S. Zhong and N. Seliya, 2005. Intrusion detection in wireless networks using clustering techniques with expert analysis. Proceedings of the 4th International Conference Machine Learning and Applications,

- [19] 10.1109/ICMLA.2005.43
- [20] Mehdi, M., S. Zair, A. Anou and M. Bensebti, 2007.
  A Bayesian networks in intrusion detection systems.
  J. Comput. Sci., 3: 259-265. DOI: 10.3844/jcssp.2007.259.265
- [21] Nakayama, H., S. Kurosawa, A. Jamalipour, Y. Nemoto and N. Kato, 2009. A dynamic anomaly detection scheme for AODV-based mobile ad hoc networks. IEEE Trans. Vehicular Technol. 58: 2471-2481. DOI: 10.1109/TVT.2008.2010049
- [22] Ohsita, Y.,T. Miyamura, S. Arakawa, S. Ata and E. Oki *et al.*, 2007. Gradually reconfiguring virtual network topologies based on estimated traffic matrices. Proceedings of the 26th IEEE International Conference on Computer Communications, May 6-12, IEEE Xplore Press, Anchorage, pp: 2511-2515. DOI: 10.1109/INFCOM.2007.311
- [23] Syurahbil N. Ahmad, M.F. Zolkipli and A.N. Abdalla, 2009. Intrusion preventing system using intrusion detection system decision tree data mining. Am. J. Eng. Applied Sci., 2: 721-725. DOI: 10.3844/ajeassp.2009.721.725
- [24] Seamans, J.K. and C.R. Yang, 2004. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog. Neurobiol., 74: 1-58.
  DOI:

10.1016/J.PNEUROBIO.2004.05.006

- [25] Tague, P., S. Nabar, J.A. Ritcey and R. Poovendran, 2011. Jamming-Aware traffic allocation for multiplepath routing using portfolio selection. IEEE/ACM Trans. Network., 19: 184-194. DOI: 10.1109/TNET.2010.2057515
- [26] Wang, Y., X. Wang, B. Xie, D. Wang and D.P. Agrawal, 2008. Intrusion detection in homogeneous and heterogeneous wireless sensor networks. IEEE Trans. Mobile Comput., 7: 698-711. DOI: 10.1109/TMC.2008.19
- [27] Yu, M. and M. Zhou, 2010. A performance modeling scheme for multistage switch networks with phasetype and bursty traffic. IEEE/ACM Trans. Network., 18: 1091-1104. DOI: 10.1109/TNET.2009.2036437
- [28] Zhong, S., T.M. Khoshgoftaar and S.V. Nath, 2005. A clustering approach to wireless network intrusion detection. Proceedings of the 17th IEEE International Conference Tools with Artificial Intelligence, Nov. 16-16, IEEE Xplore Press, Hong Kong, pp: 196-196. DOI: 10.1109/ICTAI.2005.5



**Mr.M.Thangavel** is working as Assistant Professor of Department of Computer Applications at Erode Sengunthar Engineering College, Erode, Tamilnadu, INDIA. He has received M.C.A., M.Phil Degree and currently pursuing Ph.D. at Anna University, Chennai under the guidance of Dr.P.Thangaraj, Dean, School of Computer Technologies, Kongu Engineering College. His main research interest is intrusion detection in networks. He has published one paper in National Journal, 11 papers in National Conference and 6 papers in International Conference.



Dr.P.Thangaraj received his Ph.D. degree in mathematics from the Bharathiar University, Coimbatore, India. He is currently working as Dean at School of Computer Technology & Applications, Kongu Engineering College, Erode, Tamilnadu, INDIA, has published number of papers in the National and International Journals and also published number of papers in National the and International Conferences. He is guiding more than 20 PhD scholars. Where he

teaches, among other things, fuzzy models and its usage in computer networks. He has been involved in the organization of a number of conferences and other courses at KEC, Erode. His main research interests are in Adhoc Networks of computer networks, neural networks, and pattern recognition