SIPS: A Stateful and Flow-Based Intrusion Prevention System for Email Applications

Bo-Chao Cheng*, Ming-Jen Chen*[1], Yuan-Sun Chu*,
Andrew Chen*, Sujadi Yap*[2], and Kuo-Pao Fan**
*Dept. of Electronic Engineering, National Chung-Cheng University, Taiwan
**Industrial Technology Research Institute of Taiwan
[1] m9037@cn.ee.ccu.edu.tw, [2] 94cheng@vlsi.ee.ccu.edu.tw

Abstract. In the fast-growing internet applications, email becomes more and more important in communication. SMTP attacks and spam have become one of the most serious problems. Particularly, the SMTP attacks and spam varies on email, for example spoofing address, illegal characters, sending in bulk, too many SMTP commands and so on. A single security technique is not enough to protect the system from these attacks and spam. In this paper, we propose a SMTP Intrusion Prevention System (SIPS) which bases on the concept of Stateful Protocol Anomaly Detection and Flow-based Inspection. SIPS is implemented by a finite state machine to inspect all coming email flows. It is according to the media type of email flow and their characteristics. On the test of a real email environment, our approach can prevent attacks on SMTP attack (mail bomb) average about 95.4% and spam average about 91.1%.

Keywords: Network Security, SMTP, SPAM, IPS, IDS, SIPS.

1 Introduction

In recent years, email has become more important in communication for most users on the internet. Due to the popularity and the importance of emails, many attackers try to launch SMTP attacks and spam. These problems often bother email users and administrators. Although some prevention techniques against SMTP attacks and spam are proposed respectively, these approaches usually focus on single threat. An integrated security technique is needed to resist these problems. According to the report by industry analyst firm IDC, the mark trend is changed from stand-alone threat management to Unified Threat Management (UTM). A stand-alone threat management is not enough to prevent more sophisticated email attacks. A UTM should include many security functionalities such as firewall, intrusion detection, anti-spam and so on. A robust Intrusion Prevention System (IPS) for email application must detect and prevent email attacks (SMTP attacks and spam).

SMTP attacks have become the top ten of internet security threat. Some various attacks on SMTP protocol and characteristics of spam flow [10] are as following: Buffer Overrun, Partial Message Attack, Probing Behaviors, Email Bombs, and HELO commands DoS attacks.

In order to effectively prevent both SMTP attacks and spam, we propose an integrated approach called SMTP Intrusion Prevention System (SIPS) which bases on the concept of Protocol Anomaly Detection (PAD) [2][3] and integrated with flow-based inspection to examine whether the email flows deviate from normal behavior.

The remaining of this paper is organized as following. Section 2 introduces the conventional anti-spam solutions and compares these solutions. Section 3 and section 4 describe the design concept and system architecture. In section 5, a test environment is built to evaluate the SIPS approach. And conclusion is given in section 6.

2 Background and Related Work

SMTP attacks and spam have become one of the most serious problems. In this section, we will introduce conventional anti-spam approaches and conventional prevention approaches of SMTP attacks.

2.1 Conventional Prevention Approaches of SMTP Attacks

Most of mail systems are vulnerable due to the openness of email standards and wide security holes. The attacks are including DoS (Denial of Service), buffer flow and so on. Many techniques have been developed in order to prevent such attacks. These techniques can be classified into two kinds of approaches:

A Signature-based Detection is commonly referred to the negative approach because it aims at the behavior known as “abnormal behavior” and assumes everything else as “normal behavior”.

B Protocol Anomaly Detection is commonly referred to the positive approach because it detects “normal behavior” of the specific protocol. Protocol Anomaly Detection aims at protocol misusage.

2.2 Conventional Anti-spam Approaches

Conventional anti-spam approaches can be classified into three approaches [1]:

A Content-scanning approach Saito[6] identifies email messages by analyzing mail headers and contents using keyword matching or statistical analysis of the words to determine whether an email is spam. This technique could also be called spam filtering, such as Bayesian filtering and heuristic engines.

B List-based approach [8] determines a spam by inspecting the particular IP address or email address during mail transaction before the mail server receives the email. It requires DNS to lookup the database of IP addresses that are known to be a spam source.

C Flow-based approach Qiu[7] detects spam and abnormal email behaviors in the network according to the type of email flows and their characteristics.

Qiu’s[7] flow-based concept is effective for email flows with small mail body or large amount of recipient, but it is unable to differentiate exterior and interior email behavior. Saito[6] sets a threshold according to IP address, Mail header and Mail

Body information. Spammer can deceive the system by modifying the mail header and body so that the system will not detect it (False Negative). On the other hand normal email may also be categorized as spam if it is above threshold (False Positive).

These approaches only aim at a single threat and do not have an integrated approach to prevent email attacks. We propose an integrated approach with stateful and flow-based inspection to prevent email attacks.

3 SMTP Behavior Analysis

Firstly we define the email flow and the record of an email flow for designing an integrated approach for preventing email attacks. And according to the analysis, we define the normal behaviors for an email flow.

3.1 Definitions of The Email Flow and The Record of Email Flow

An email flow is constructed of five tuple (source and destination IP, source and destination port and protocol type). As shown in Table 1, the flow record [9] is used to record behaviors of email flow (SMTP flow). It is for single direction from SMTP client to SMTP server and it holds values of attributes which interest in this flow.

According to the record of the email flow, we can analyze the behavior of an email between SMTP client and SMTP server. Based on the behavior observation, we can verify the normal or abnormal email with our finite state machine which is described in the following.

Table 1. An example for record of an Email Flow.

Filed Name	Description
FlowID	ID number of each SMTP flow.
StartTime	SMTP Connection established time.
LastTime	The newest received network packet time of SMTP flow.
TotalFlowpkts	Total outbound network packet numbers.
ConnectionInfo	Source IP, Destination IP, Source Port, Destination Port, Protocol Type
TotalFlowsizes	Total outbound packet size.
BDFlowpkts	Outbound network packets numbers after receiving DATA command.
BDFlowsize	Outbound network packets size after receiving DATA command.
HeaderSizes	SMTP flow mail header size.
BodySizes	SMTP flow mail body size.
RcptCounts	Mail Recipients' number in a SMTP flow.
PlainFlag	Mail type in plain Mail.
HtmlFlag	Mail type in html Mail.
EmbeddedFlag	Mail with embedded resource.
AttachmentFlag	Mail with attachment files.
FormFlag	Mail with html form.
OutSideFlag	SMTP Client's IP address is Interior or Exterior.

3.2 Normal Email Behavior

Firstly we observe the behavior of a normal email by monitoring more than 300 emails transferred to a SMTP server. We specify that a normal email flow should have the following characteristics:

- In general, there are many differences between some MTA software, but a normal mail behavior usually follows RFC formats such as minimum implementation, general syntax principles and transaction model.
- A transmission of normal mail should not spoof the domain of SMTP client and make a guess on username.
- The same IP address of SMTP client sending emails should be in the regular cycle.
- The variation of the email flow size depends on the behavior of email flow mentioned above, the location of SMTP client and the media type of emails.
- An email flow that contains a mail transaction must be initialized by using EHLO command. A mail transaction includes several SMTP commands, which are MAIL, RCPT, DATA and QUIT. The SMTP client should send these commands to the SMTP server in order to expect QUIT command.
- By RFC2821 [5], the NOOP, HELP and VRFY commands can be used at any time during an email flow, but a normal mail mostly does not contain these SMTP commands.

Based on the normal email behavior, we can now design the SMTP Intrusion Prevention System based on stateful and flow-based inspection.

4 SMTP Intrusion Prevention System (SIPS)

Intrusion Detection System (IDS) has become an important part of network security system on most of business enterprises. Its function is like security alarm or surveillance system in our houses. If an intrusion is detected, it will alert the network administrator for further process. IDS could only passively detect an intrusion, but could not prevent it. It is insufficient for enterprises network requirement. Therefore a new generation of improved system is proposed, Intrusion Prevention System (IPS) with a feature to prevent an intrusion. In this paper we proposed a SMTP Intrusion Prevention System (SIPS) which can protect email servers from attacks.

4.1 System Architectures

We integrated SIPS with Snort[4] (Snort is an open source Network-based IDS), to aim at email attacks based on SMTP behaviors of practice and study of RFCs.

As shown Fig.1, a new preprocessor called “SMTP Inspection Preprocessor (SIP)” is implemented and integrated with Snort to realize our approach. The decoder resolves the protocol which is used by the given packet and matches the data against allowable behavior for packet of their protocol. After the packets are matched, the preprocessor will redirect the email flow to SIP, inspecting whether they are email attacks. And then the email flow will be taken into Detection Engine and compare it against the rules in Snort without verdict in SIP. It is the core part on signature-based NIDS.

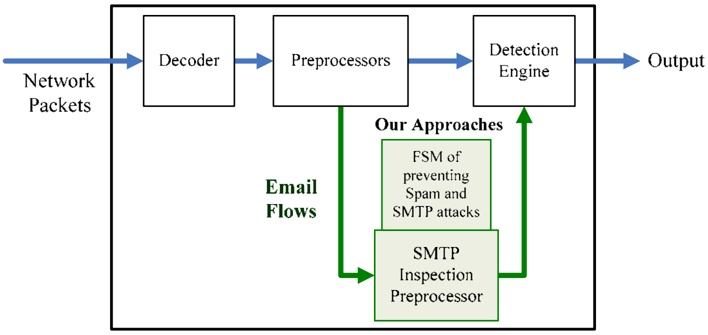

Fig. 1. SMTP Intrusion Prevention System, SIPS

4.2 SMTP Inspection Preprocessor (SIP)

SIP is constructed of two components as shown in Fig. 2. The first component is a finite state machine (FSM) for preventing email attacks whose states are mainly from RFC2821. It allows us to detect and prevent any deviation from the normal email behaviors we specified. The other component is the anomaly behaviors of email flows. It gives a feedback for administrators so that the administrators can re-specify the detection variables to reduce false positive.

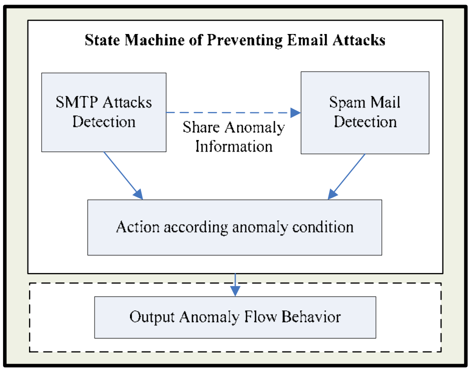
State Machine of Preventing Email Attacks
SMTP Attacks
Detection
Share Anomaly
Information
Spam Mail
Detection
Action according anomaly condition
Output Anomaly Flow Behavior

Fig. 2. SMTP Inspection Preprocessor

Protocol processes can be modeled as a collection of communicating FSMs. FSM of protocols must store some data value and synchronization messages [9] to maintain the temporal order of the event. In this paper, the SMTP communication behaviors are modeled as FSM based on the email flow. In order to check sizes, length, syntax and the order of SMTP commands and its parameters, each command is modeled as a state. It is constructed in a Moore machine, so that its output depends on the state and the input.

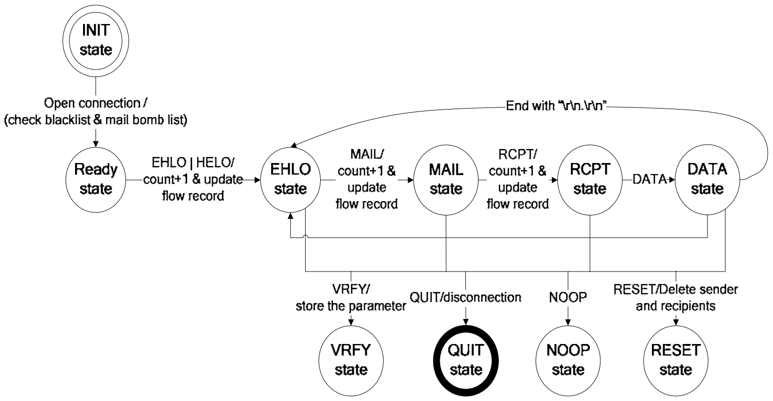
INIT state
Open connection / (check blacklist & mail bomb list)
Ready state
EHLO | HELO/ count+1 & update flow record
EHLO state
MAIL/ count+1 & update flow record
MAIL state
RCPT/ count+1 & update flow record
RCPT state
DATA
DATA state
End with "\r\n.\r\n"
VRFY/ store the parameter
QUIT/disconnection
NOOP
RESET/Delete sender and recipients
VRFY state
QUIT state
NOOP state
RESET state

Fig. 3. The State Machine of Preventing Email Attacks

Fig. 3 shows the finite state machine of SIPS, the state transitions and prevention approach are described as following:

- Initial State: As SMTP connection is established, the system enters this state and checks whether the IP address is on blacklist and mail bomb list. The connection is categorized as normal if it is not on the list. The flow information will be stored and updated to flow records which are used to determine SMTP attack and spam.
- Ready State: After TCP 3-way handshake, the system enters this state that is waiting for HELO or EHLO command sent by SMTP Client. On this state the system sets a period of time Ta. The email flow is considered as normal if the number of connection is below the threshold N during Ta. Otherwise it is likely to be an email bomb, the system discard the email flow and record the source IP address in the mail bomb list for a dynamic and real-time prevention.
- EHLO State: According to RFC2821, the variable of HELO command must be a Full Qualified Domain Name (FQDN). In this state the SMTP client must send HELO/EHLO command with variable which follows FQDN. The length is limited under MaxCommandLength to prevent buffer overflow.
- MAIL State: The sender address must follow "<sender@domain>" format. And this domain must follow FQDN and checked for its existence. Attacker and spammer usually use false domain name to deceive SMTP server.
- RCPT State: The recipient address must follow "<recipients@domain>" format. The SMTP server will authenticate whether the sender or recipients is a legal user. If it is an illegal user, SMTP flow will be delayed 20 seconds, and if another illegal user is received, it will be delayed for another 30 seconds. If it keeps receiving illegal users, then this SMTP flow is considered to be abnormal.
- RESET State: The system enters this state if the previous mail transaction is canceled and the sender and recipients will be deleted. This flow is considered to be abnormal if the command is larger than 6 byte.
- VRFY State: If the number of failed user verification or VRFY command is sent more than the threshold, this SMTP flow is considered to be abnormal.

- QUIT State: Entering this state, the timer of this flow record can be stopped. This flow is considered to be abnormal if the command is larger than 6 byte.
- NOOP State: This command will not affect any previous command, except an OK is sent if this command is received.
- DATA State: Entering this state means that envelop commands (“Mail From” and “Rcpt To”) are done. Then the SMTP clients will send DATA commands. And after the SMTP server replies with code message 354, it will begin to receive the mail content from SMTP client. An email flow is considered as normal if the DATA command sent is below the threshold. SMTP commands order must be followed to enter the DATA State, otherwise the email flow is considered to be an abnormal flow.

The behavior of each email flows will be analyzed with the SIPS’ finite state machine in each state transition. And the SMTP attacks and spam will be prevented. In the next section, we will show our experimental results.

5 Experimental Results

We construct a real test environment shown in Fig. 4 to evaluate our approach.

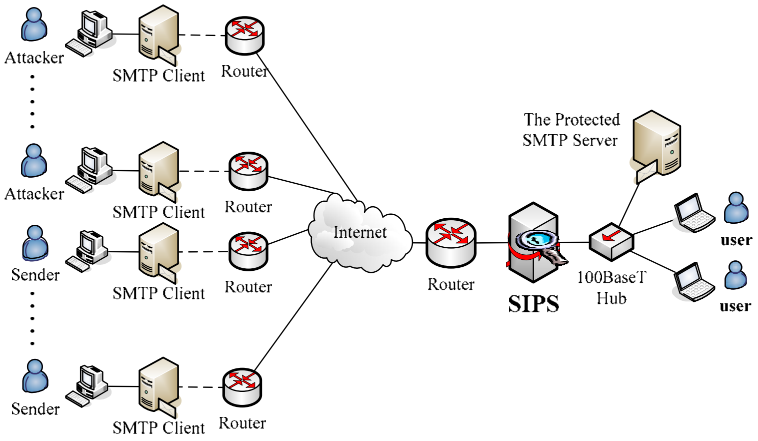

Fig. 4. Test Environment

On the left hand side, the senders (including normal sender, attacker and spammer) send emails to the Protected SMTP server on the right hand side. The SMTP Server on the right hand side is protected by our SMTP Intrusion Prevention System (SIPS) which uses stateful and flow-based method to detect SMTP attacks and spam. We evaluate performance of SIPS in term of false positive (FP) and false negative (FN) [14]. The attackers attack the SMTP server protected by SIPS through the SMTP client with open relay or free web mail service. And some sender sends normal email. Emails need to be inspected before sent to the protected SMTP server. In our testing, we use email bombs and spam to evaluate our approach correctly and efficiently. In evaluating spam, we compare SIPS with two papers Saito[6] and Qiu[7] mentioned in section 2.

Test of Mail Bomb

As shown in Table 2, we send 50, 100, 150, 200, 250, 300, 350 mail bombs with two IP addresses and 100 normal mails as the background traffic.

(100,50) denotes 100 normal mails and 50 mail bombs. As shown in Fig.5, the ratio of FP is always below 5%. It is affected with a main factor: normal users might send a large number of mails and cause SIPS producing a false alarm. SIPS is still better than the existing security technology and the original Snort which can not detect and prevent any attacks of mail bombs. The ratio of FN is always below 10%. In “INIT state” and “Ready state”, the FSM can detect and prevent the email bombs. FN of our approach is converged as the email bombs grow.

Table 2. Mail Bomb Test Traffic

Case	Normal Mails	Mail Bombs	Total
A	100	50	150
B	100	100	200
C	100	150	250
D	100	200	300
E	100	250	350
F	100	300	400
G	100	350	450

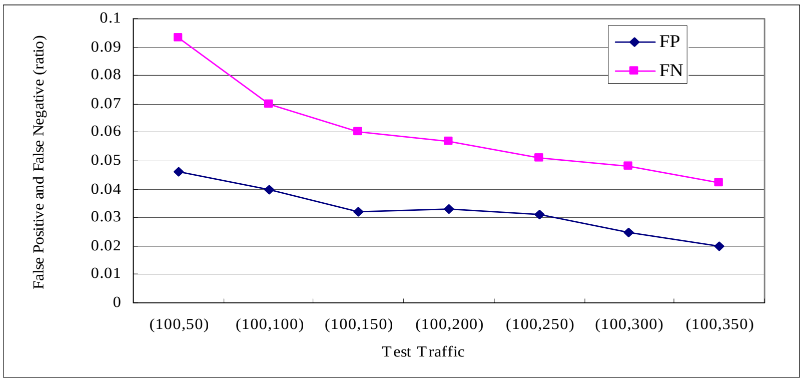

Fig. 5. SIPS Results on Mail Bombs Prevention

Test of Spam Mails.

Our proposed approach can prevent spam according to normal behavior constructed by finite state machine. The results are shown in Table 3, Fig. 6, and Fig. 7.

As shown in Fig. 6, the FP of the approach proposed by Qiu[7] is worse than our approach, because it only determines spam by the flow size, there is no difference regarding interior or exterior email server. The same inspection for interior and exterior server makes the FP increasing. Qiu[7] approach also can not detect spoofing and defrauding during mail transactions.

The approach proposed by Saito[6] determine the email is spam when it receives large number of same mail header or mail body, or a large number of email from the same IP address in a period of time. But some of normal emails also have these characteristic so that it may lead to worse FP.

Table 3. Comparison of Preventing Spam

Normal	Spam	FP-SIPS	FN-SIPS	FP-Qiu	FN-Qiu	FP-Satio	FN-Satio
100	50	10.6%	5.3%	16%	8.7%	22%	14%
100	100	9%	9%	12%	12.5%	16%	18%
100	150	6.4%	11.6%	10.4%	14%	12%	18.4%
100	200	5.3%	12.3%	9%	14%	12%	18.6%
100	250	5.7%	12.9%	6.6%	14%	10%	18.9%
100	300	4%	14.3%	7.8%	15.8%	8.8%	18.8%
100	350	3.6%	15.1%	6.9%	16.7%	6.9%	18.4%
Average		6.3%	11.5%	9.8%	13.6%	12.5%	17.8%

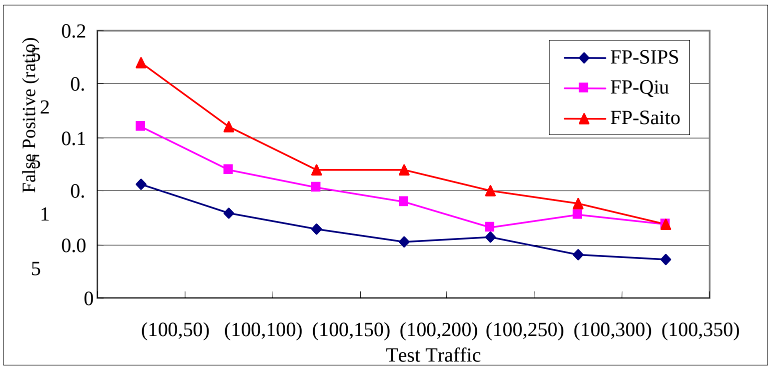

Fig. 6. Comparisons of Preventing Spam (FP)

As shown in Fig. 7 FN of the approach proposed by Qiu[7] is worse than our approach. It is because this approach does not record the state information. It only detects spam by quantity of the flow. Spammers usually change their identification within mail transaction so that their real identification will not be discovered.

Saito[6] approach is also insufficient to prevent spam because it uses IP header information. Nowadays, spammer usually sends spam by changing IP (SMTP client), header and body randomly. If spam has these characteristics, this approach will have worse FN.

SIPS uses stateful deep inspection on each email flow targeting various type of spam. The results show that SIPS is better than booth Saito[6] and Qiu[7] for the test traffic. The average ratio of FP in SIPS is 6.3% and 11.5% for the FN.

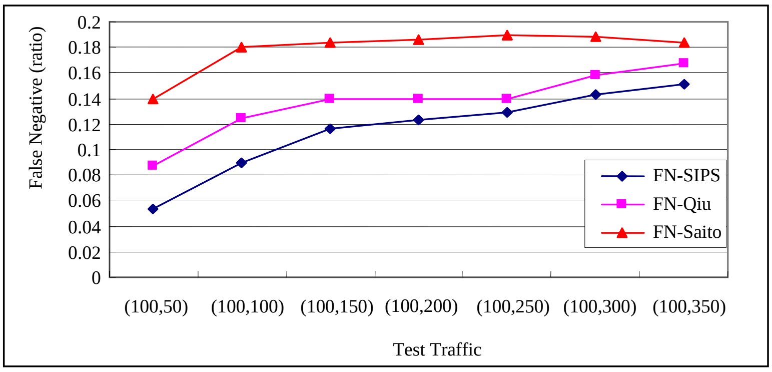

Fig. 7. Comparisons of Preventing Spam (FN)

6 Conclusions

Since 1982 SMTP has become one of the most important internet applications, it has also become attacker's target. This paper proposes an approach to prevent spam on mail transactions using stateful and flow-based inspection on email flows. To blend a series of security solutions we also integrate Snort into IPS for email application. We evaluate our approach with real email environment. The average results are 3.2% for FP, 6% FN on mail bomb test and 6.3% for FP, 11.5% FN on spam mail test.

References

[1] D. Harris, "Drowning in Sewage: SPAM, the course of the new millennium: an overview and white paper," SpamHelp, Available: http://www.spamhelp.org/articles/Drowning-in-sewage.pdf

[2] K. Das, "Protocol Anomaly Detection for Network-based Intrusion Detection," SANS, August 13, 2001

[3] CERT, "State of the Practice of Intrusion Detection Technologies". Available: http://www.cert.org/archive/pdf/99tr028.pdf

[4] Snort, Available: http://www.snort.org

[5] J. Klensin, "SIMPLE MAIL TRANSPORT PROTOCOL," RFC2821, April 2001

[6] T. Saito, "Anti-SPAM System: Another Way of Preventing SPAM," Database and Expert System Applications, 2005. Proceedings. Sixteenth International Workshop on 22-66 Aug. 2005 Page(s): 57 - 61

[7] X. Qiu, J. Hao, & M. Chen, "Flow-based anti-spam," IP Operations and Management, 2004. Proceedings IEEE Workshop on 11-13 Oct. 2004 Page(s):99 – 103.

[8] J. S. Park, & A. Deshpande, "Spam Detection: Increasing Accuracy with a Hybrid Solution", Information System Management, winter 2006

[9] N. Brownlee, "Traffic Flow Measurement: Architecture", RFC2722, October 1999

[10] T. Bass, & A. Freyre, "E-mail bombs and countermeasures: cyber attacks on availability and brand integrity", Network IEEE, 1998