A Safety Quota-variable Mail Transfer Protocol under Isolated Environment

Jiangli Zhang
Guangdong Polytechnic of Environmental Protection Engineering, Foshan, China, 528216
Email address: jiangli@just.edu.cn

Abstract-The information security problems of e-mail have become a major threat of the network information security in recent years. In this paper, we did deeply analyze on the principles and safety of the protocols of e-mail system, such as SMTP and ESMTP. An improved protocol based on SMTP and ESMTP was proposed, and protocol mode was designed as well. The protocol was denominated as Safety quota-variable Mail Transfer Protocol (SVMTP) under isolated environment. The results show that svmtp protocol has more obvious effect in blocking suspicious connections, security classification and minimizing the waste of communication resources.

1. INTRODUCTION
The main technical reasons for the information security problems of e-mail are: ineffective mail sending control measures, simple blocking techniques for unsafe mail, and passive virus prevention. The current anti-spam technologies and other related measures cannot effectively control the proliferation of spam. Based on an in-depth analysis of the technical reasons for spam generation and development, this research improves the mail protocol and organically integrates multiple technologies to form a multi-level and multi-layer interception system and effectively improve the status quo of spam. The main work of this paper is to conduct an in-depth analysis of the SMTP transmission process and working principle. Aiming at the security limitations of the SMTP protocol, an improved protocol SVMTP based on the SMTP protocol is proposed. A new mail protocol SVMTP is designed and analyzed. Three security modules were added on the basis. Finally, the advanced nature of the program was discussed with a comparative analysis method.

2. Analysis of SMTP E-mail Transmission Process and Working Principle
SMTP (Simple Mail Transfer Protocol) is a member of the TCP/IP protocol suite. It is a set of rules used to transfer mail from a source address to a destination address. It controls the way of mail transfer and the receiving/transfer of emails.

2.1 SMTP model and command set
A two-way transmission channel is established between the sending SMTP server and the receiving SMTP server, as shown in Figure 1. In the communication process, the sender SMTP and the recipient SMTP adopt a conversational interaction mode. The entire process is controlled by the sender and sometimes requires several confirmations. In order to ensure the effectiveness of the reply command, SMTP requires the sender to provide the recipient's server and mailbox. Both sender and recipient have strict grammatical definitions and corresponding digital codes in the form of ASCII codes. The
SMTP command defines the mail transmission function. The data shows that command are completely ignored for the safe transmission process and process control, because the operation of sending mail involves different data objects or parameters.

![Communication connection flow chart between servers](image)

**Fig. 1. Communication connection flow chart between servers**

### 2.2 SMTP protocol security analysis and ESMTP

The SMTP protocol allows arbitrary declaration of mail source information, no verification of the authenticity of the source identity, and no judgment on the authenticity of the address protocol. ESMTP (Extended Simple Mail Transfer Protocol) is an extension of SMTP, which is slightly stronger than SMTP in terms of security. The protocol provides an authentication module for the MTA host. The AUTH command is introduced [1] [2], and the client host provides identity authentication to the mail server by using the AUTH command. Providing identity authentication can identify the authenticity of the source of the mail, but it cannot accurately determine the cause of the false mail, and cannot solve the behavioral problem of sending spam. [6] [10]. Therefore, it is very important to establish a mail transmission protocol that can recognize and block spam from multiple angles through comprehensive means. This research proposes the E-mail Security Extension Protocol (SVMTP).

### 3. SVMTP

#### 3.1 Design ideas of SVMTP

The improvement of SVMTP in mail security has three modules: the MTA authentication module to solve the false source problem, the MTA isolation classification module to solve the classification process, and the connection quota variable module to solve the large-scale automatic forwarding problem. The steps are as follows:

**Step1: Set feature threshold classification processing:** Compare the mail feature weighted index M obtained from the judgment with the system's spam threshold index to determine the nature of the mail. The results are: normal mail, spam and suspicious mail.

**Step2: Restrict the automatic forwarding behavior mechanism:** The virus software propagation characteristic is that as long as the user opens the mail, the virus program will start to automatically search the mailing list address and forward it. Unable to ensure the user's information security and the loss of network resources occurred. The ideal solution is to be filtered in the first forwarding MTA server.

**Step3: Control artificial group distribution measures:** Technical restrictions on group distribution can restrict spammers from sending a large number of e-mails in a short period of time, and set rules for communication connection frequency. Set the number of connections attempted from
the same source address within a period of time to control a reasonable preset range, thereby blocking the connection of the spam server.

3.2 The overall design of SVMTP model
The SVMTP model is a security protocol extension based on ESMTP. The model deals with the security of three aspects: false source, classification, and large-scale forwarding of mail from the original source address. The model diagram of SVMTP is as follows:

![SVMTP security module structure diagram](image)

When sending the e-mail request, if it comes from the local, as shown in the two authentication steps shown in Figure 1 and 2, the authentication result will be sent back to the MUA. If the authentication is passed, the process 3 sends a ready command to the classification and isolation module of the MTA. The spam classification and isolation module has multi-layer classification processing capability. The ready classification and isolation module accepts process 4 and starts sorting. According to the processing result of the spam content characteristic factors, the mail is put into one of the normal mail queue, the spam queue and the suspicious mail queue.

3.3 Isolation and classification process
The isolation processing module is divided into three layers to carry out classification work:

The first layer: Separation of spam based on the characteristics of the network layer mail. The end of this level of classification means that the system quickly and accurately processes a large number of spams, thereby reducing the load on the server for further work.

The second layer: Judging the nature of the mail, and weighting the spam characteristic factors of each layer. The weighted judgment index is compared with the reference index to judge the nature of the mail.

The third layer: Filter classifier, with the help of processing technology in this paper to achieve the third processing of suspected spam.

3.4 Multi-feature classification process of isolation classifier
First, obtain a comparison index $\rho$ of this layer through the judgment experience of a large number of spam emails, and then form a judgment index $M$ for the weighted sum of the characteristics of
each suspected email sent from the upper layer. Compare this judgment index $M$ and the existing comparison index $\rho$ in the layer. The introduction of $\delta$ re-divides the traditional dichotomous results into three intervals, representing three types of mail. The value of $\delta$ comes from the mean square error in the system statistics process:

$$\delta = \sqrt{\frac{(M_1 - \rho)^2 + (M_2 - \rho)^2 + \ldots + (M_n - \rho)^2}{n}}$$

(1)

In this formula, the $M_i$ value comes from the weighted sum calculated by the system for the $i$-th mail object, and the $\rho$ value is the statistical average value of the mail system. Only the $M$ value calculated for the current mail is called the judgment index. The calculation formula of $M$ is:

$$M = \sum_{i=1}^{N} a_i C_i$$

(2)

In this formula, $C_i$ is a feature, if the feature exists, it is 1, and if the feature does not exist, it is 0. $a_i$ is the feature weighting coefficient. $N$ is the characteristic number of spam which is got from analyzing by the current judgment layer of the mail. The value of $a_i$ and $C_i$ is a feature parameter table defined according to the importance of features to judgment. The features in the spam feature data table are derived from the statistics of a large number of spam samples. The spam samples are mainly derived from spam reports from email users to provide real-time updates. The weights of judging spam in the system are shown in Table 1:

**TABLE 1 Feature parameters and weight tables involved in judging and classifying spam at the network layer**

| Number | Spam characteristics                        | C value | Eigenvalue weight(a) |
|--------|---------------------------------------------|---------|----------------------|
| 1      | X-mailer field no                           | 1       | 0.5                  |
| 2      | X-mailer field is special                   | 1       | 0.7                  |
| 3      | Different from the From field               | 1       | 0.9                  |
| 4      | Reply-to reverse analysis exception         | 1       | 1                    |
| 5      | From field contains keywords               | 1       | 0.8                  |
| 6      | RCPT command address is inconsistent with   | 1       | 0.9                  |

When an email arrives, the initial value of $M$ is 0. Whether an email is spam or not is determined by comparing the $M$ value obtained with the comparison index $\rho$. The judgment formula is as follows:

$$\begin{align*}
\text{Mail classification basis} & \begin{cases}
\text{Spam} & M > \rho + \delta \\
\text{Normal mail} & M < \rho + \delta \\
\text{Suspected mail} & \text{other}
\end{cases}
\end{align*}$$

(3)

The entire isolation classification process is shown in Figure 3:
3.5 Traffic customization and variable quota control over spam

The behavior characterized of spam is high connection speed and frequency. The separated mails, regardless of their nature, must be distributed with rules to restrict large-scale distribution. SVMTP distribution rules are as follows:

Rule 1: In order to avoid missed mail inspection, in this protocol, even if the normal mail EMSG1 is not directly forwarded, it must be processed through the flow control module. The flow limit valve unit is that the number of connections requested by the current mail to be forwarded within a certain period of time T is within the specified range of the system, and the current mail of the system opens a customized maximum connection value in a normal category.

Rule 2: The suspected mail EMSG2=(fromip2,rcpt2, size2,linkmsg2) is a four-tuple. EMSG2 defines the relationship between the evaluation factor of suspected mail and the characteristics of the mail, where fromip2 is whether the mail comes from the local area. If it is a local mail, the value of fromip2 is 0, otherwise it is 1. When the item of fromip2 is 1, it is a normal forwarding condition. Rept2 refers to the number of recipients and it is an integer greater than zero. When the value of Rcpt2 is 1, it will be forwarded normally. When the value of Rcpt2 is greater than 1 and less than or equal to a limit value Max customized by the system, Max is a statistical average value of the RCP field information of this mail system. If the first two tuples pass, check the mail size (size2) and the number of connections (linkmsg2). If the latter two values are within the forward able range, and both size2 and linkmsg2 are 0, then set Rcpt2 to 1, which means normal forwarding. But if size2 and linkmsg2 is not within the forward able range, the height of the valve is lowered and the number of forwarding mails is limited, so as to effectively curb the mass distribution of spam. The processing flow of suspected mail EMSG2 is based on the normal mail EMSG1, adding two layers of judgment to ensure that the system handles the suspected mail more accurately.

4. Analysis and Evaluation of SVMTP Scheme

The SVMTP protocol solves some problems existing in the current mail system protocol in terms of security, and has certain advantages over SMTP and ESMTP and some existing improved application technologies.

4.1 Comparison of SVMTP and ESMTP

ESMTP is an extension of the authentication of the SMTP protocol in terms of mail security. The basic idea is that the server rejects issues such as forged source addresses and false mail routing during the mail sending process [5] [10]. The solution only solves the problem of false sources of spam. SVMTP is designed with three layers of security measures: the authenticate on subsystem uses the
authentication technology provided by ESMTP to determine whether the source is false. The isolation sub-module determines the classification to isolate the spam queues with obvious characteristics. Normal mail enters the behavior control module, and the flow control algorithm is activated to determine the number of connections. The suspected type queue is weighted and summed to obtain the judgment factor according to the characteristic table, and the judgment factor is compared with the comparison factor to complete the secondary classification work. A small number of suspected emails after two classifications entered the third layer of content analysis and detection. Multi-layer classification control is designed on false source, content and connection behavior.

4.2 SVMTP filtering VS traditional filtering

Traditional filtering technology simply uses filtering technology. Email security problems still exist. The filtering process is overly dependent on the database. Spammers quickly adapt to the filter after dynamic learning and bypass the filter. New unsafe emails will pass through the server which will produce misjudgment and omission. The SVMTP protocol has obvious advantages in terms of security. First, it has the characteristics of multi-layer classification based on spam characteristics. Second, there is a separation module for judging spam based on comprehensive characteristics. Besides, the system has a spam feature table with self-learning ability. What’s more, the judgment module adds suspicion classification.

5. Conclusion

There are obvious security risks in the security design of the current mail protocol. The technical limitations of email security technology for a long time have not controlled the proliferation of spam. Anti-spam technology cannot be effective by simply identifying quarantine, blocking and deleting. There are several types of mail from the separation module. In addition to the spam queues with obvious characteristics which are quarantined, and normal mails with low connection requests without spam characteristics which are forwarded, there are suspected mails and normal mails with large-volume connection requests. SVMTP uses flow customization technology and variable quota technology to control the number of spam that has not been quarantined successfully. The flow control technology is like adding a throttle to the data flow before reaching the forwarding queue of the email server. When encountering suspected spam, the throttle valve will be narrowed and the variable quota scheme of the throttle valve will be activated. The control of a certain number of connections is dynamically adjusted by the parameters brought by the previous separation. The improved protocol has obvious effects on spam processing and has certain theoretical significance.

REFERENCES

[1] Mohamed F A, Hemayed E E. Using trusted computing in trusted mail transfer protocol[J]. Security & Communication Networks, 2014, 7(5):926-933.
[2] Enciso L, Baez R, Maldonado A, et al. E-Mail Client Multiplatform for the Transfer of Information Using the SMTP Java Protocol Without Access to a Browser[M]// Trends and Advances in Information Systems and Technologies. Springer, Cham, 2018.
[3] Mi B, Huang D, Wan S, et al. A post-quantum light weight 1-out-n oblivious transfer protocol[J]. Computers & Electrical Engineering, 2019, 75:90-100.
[4] Lee C C, Chen S D, Li C T, et al. Security enhancement on an RFID ownership transfer protocol based on cloud[J]. Future Generation Computer Systems, 2018.
[5] Dijk B V, Meijer J, Cuypers T D, et al. Trusting the hand that feeds: microbes evolve to anticipate a serial transfer protocol as individuals or collectives[J]. BMC Evolutionary Biology, 2019, 19(1):201-205.
[6] Xie R, Jian B Y, Liu D W. An improved ownership transfers for RFID protocol[J]. International Journal of Network Security, 2018, 20(1):149-156.
[7] Venkauskas A, Morkevicius N, Jukavius V, et al. An Edge-Fog Secure Self-Authenticable Data Transfer Protocol[J]. Sensors, 2019, 19(16):3612.
[8] Kim D, Baek J, Fisher P S, et al. fFTP: a fast file transfer protocol for home N-screen platform[J]. Personal and Ubiquitous Computing, 2018, 22(1):143-152.

[9] Ibrahim A, Zhao L. Supporting the OSGi Service Platform with Mobility and Service Distribution in Ubiquitous Home Environments[J]. The Computer Journal, 2018, 52(2):210-239.