# 스팸메일 차단을 위한 SMTP 보안 게이트웨이 설계

이창성*, 이은선*, 한영주*, 김희승*, 정태명*
*성균관대학교 정보 통신 공학부
이동형 응급의료 정보 시스템 개발센터
e-mail:{cslee, eslee, yjhan, hskim}@imtl.skku.ac.kr,
and tmchung@ece.skku.ac.kr

# Design of the Secured SMTP Gateway for Spam-Mail Interception

Chang-Sung Lee*, Eun-Sun Lee*, Young-Ju Han*,
Hee-Seung Kim*, and Tai-Myoung Chung*
1)*Cemi: Center for Emergency Medical Informatics
School of Information and Communication Engineering,
Sungkyunkwan University

요 약

인터넷이 발전함에 따라 기업의 업무, 커뮤니케이션 등이 온라인으로 전환되고 있으며, 정보 전달의 통로로써 전자 메일의 사용이 나날이 늘어남과 동시에 전자 메일을 통한 스팸메일의 폭발적인 증가로 인한 심각성 또한 대두되고 있다. 현재 스팸메일을 막기 위한 여러 가지 방법이 제안되었으나, 대부분 메일 서버내의 정책에 따른 메일 필터링 방식으로써 완벽한 스팸메일 탐지를 제공하지 못하며, 스팸메일로 인한 메일서버 및 네트워크 자원 손실 문제는 여전히 해결되지 않고 있다. 본 논문에서는 스팸메일 탐지율을 높이고 네트워크 내 자원 손실을 예방할 수 있는 SMTP 보안 게이트웨이를 제안하고자 한다. 본 SMTP 보안 게이트웨이는 스팸메일 차단 규칙에 의한 메일 필터링을 기본적으로 제공하고, 룰에 정의되지 않은 메일에 대해서는 사용자 선택에 기반한 메일 전송을 제공한다. 이는 규칙에 정의되지 않은 스팸메일에 대한 탐지 가능성을 높이며, 궁극적으로 메일서버의 자원 및 네트워크 자원의 가용성을 높일 수 있다.

## 1. 서론

인터넷이 급속히 확산되고 발전된 지금, 인터넷을 기본으로 하여 중요한 통신 수단으로 자리 잡은 전자 우편 서비스는 현재 '스팸메일'의 형태로 악용되는 부작용을 겪고 있는 현실이다.

스팸메일은 수신을 원하지 않는 사람에게 강제적으로 전자 우편을 보내는 것을 의미하며, 이는 최근 몇 년 동안 대두된 심각한 문제이다[4]. 스팸메일은 무차별적으로 다량 전송되기 때문에 개인 사용자의 시간적, 금전적 피해 뿐 아니라 메일서버 자원과 네트워크 자원의 손실을 가져온다. 현재 많은 메일 필터링 방식이 제안되고 있으나, 대부분이 개별 사용자의 보호에 한정되는 한계를 가진다. 즉 기존 스팸메일에 대한 대응책은 개별적인 사용자나 메일서버에 의한 정책에 의해서 이루어지고 있어, 네트워크와 메일서버의 자원 손실을 해결하지 못할 뿐만 아니라, 스팸메일의 형태가 매우 다양해지고 있어 완벽한 스팸메일 차단을 하지 못하고 있다.

본 논문에서는 게이트웨이 수준에서 내부 네트워크로 들어오는 모든 메일에 대하여 효과적으로 스팸메일을 차단할 수 있는 스팸메일 방지 게이트웨이를

본 논문은 보건복지부 보건의료기술진흥사업회 지원에 의하여 이루어진 것임(과제번호: 02-PJ3-PG6-EV08-0001)

제안한다. 이는 스팸메일 방지 게이트웨이에서 내부 네트워크로 들어오는 메일을 수집하여, 스팸메일을 분류하고, 스팸메일로 분류되지 않은 메일에 대하여 메일서버에 해당 메일에 대한 기본 정보만을 전송한 뒤 사용자의 선택에 따라 메일원문을 전송하는 시스템이다. 이와 더불어 내부 네트워크의 가용성 향상을 위해 메일서버 자원과 내부 네트워크 자원을 보호할 수 있는 방법에 대해 논의할 것이다.

본 논문의 구성은 다음과 같다. 2장에서 스팸메일에 대한 현재 연구 사례와 본 논문에서 제시하는 시스템과 관련된 연구들을 살펴보고, 3장에서는 본 논문이 제안하는 SIG 시스템(Spam-mail Interception Gateway System)의 구성요소와 각 요소별 기능 및 구조에 대해서 설명하고, 4장에서는 SIG 시스템의 장점 및 활용, 5장에서는 결론 및 향후 계획에 대하여 기술한다.

## 2. 관련 연구

### 2.1 SVM(Support Vector Machine)을 이용한 스팸메일 필터링

개별적인 사용자를 위한 스팸메일 필터링 방식인 SVM은 자기학습을 통하여 스팸메일을 분리하는 알고리즘으로 이진 분류 문제에 있어서 가장 효율적으로 알려져 있다[1]. SVM은 입력패턴들을 교사학습 방법을 통하여 +1과 -1의 두 클래스로 패턴을 분류하고, 두 개의 클래스로 분리하는 기준인 하이퍼플레인(hyperplane)이 결정한다. 하이퍼플레인을 결정하는 입력 패턴들을 support vector라 하며, 정확한 하이퍼플레인은 잘못된 분류를 막을 수 있다. 스팸메일을 탐지하기 위하여 우선 데이터 사전을 생성하게 되는데 데이터 사전은 SVM의 학습데이터로 쓰이게 되고 학습을 수행한 SVM은 스팸메일 여부를 판단하게 된다. 하지만 SVM은 스팸메일 형태의 다양성 때문에 완벽한 스팸메일 차단에는 한계가 있고, 메일서버와 네트워크의 자원을 보호하지 못한다.

### 2.2 의도적인 메일주소 유포를 이용한 스팸메일 구분 방법

이 기술은 일반적으로 스팸메일을 송신자가 인터넷 게시판 등에 공개된 메일 주소를 수집하여 주소 리스트를 만들어 사용하는 경우가 많다는 점을 착안하여 만든 것으로 데이콤이 기술을 고안하여 특허를 신청한 것이다[5]. 실제 존재하지 않는 메일 주소를 의도적으로 유포한 후 스팸메일 송신자에게 메일 주소가 유입되도록 한다. 메일서버에서 이 메일 주소가 포함된 메일은 스팸메일로 간주하여 차단하는 방법이다. 이것은 스팸메일을 알아내는 효과적인 방법이지만, 결국 메일서버는 스팸메일을 받게 되어, 네트워크와 메일서버의 자원이 낭비된다.

### 2.3 애플리케이션 게이트웨이

애플리케이션 게이트웨이는 OSI 7 계층의 애플리케이션 계층에 해당되는 데이터 영역의 패킷을 정해진 정책에 의하여 검증하고 패킷의 통과 여부를 판단하는 방식을 지원한다[2]. 즉 클라이언트에서 서비스 요청(Telnet, HTTP, FTP)이 들어오면 방화벽에서 애플리케이션 데이터 검증을 거쳐 목적 시스템으로 접속되는 방식이다. 애플리케이션 게이트웨이는 실제 데이터 확인을 통한 강력한 보안이 가능한 장점을 가지고 있다. 본 논문에서 제시하는 SIG 역시 애플리케이션 계층 데이터까지 확인하는 애플리케이션 서버의 한 종류라고 할 수 있다.

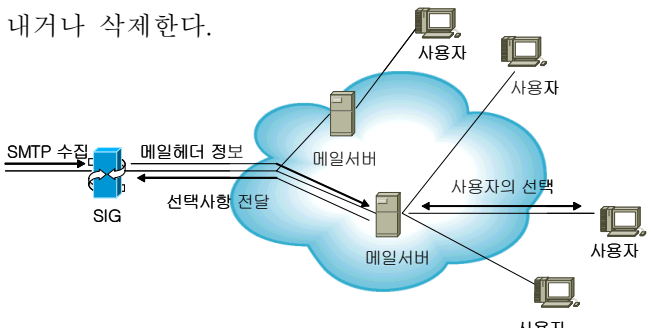
## 3. SIG 시스템

내부 네트워크와 메일서버를 스팸메일로부터 보호하기 위한 SIG 시스템은 내부 네트워크로 들어오는 SMTP 메시지를 수집하여 필터링을 수행하고 저장하기 위한 SIG, 사용자의 의사를 SIG에 전달하기 위한 SIGP(SIG Protocol), 그리고 SIGP 관련 모듈로 구성된다.

### 3.1 SIG

#### 3.1.1 SIG 기능

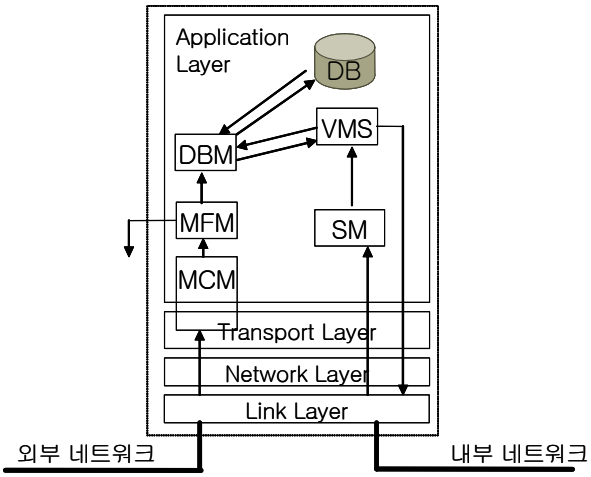
SIG는 [그림 1]과 같이 외부 네트워크와 내부 네트워크를 사이에 위치하여 이 사이를 오가는 모든 패킷 중 SMTP 관련 패킷을 흡수하고, 1차적으로 이미 정의된 스팸필터링 규칙에 의하여 필터링을 수행한다. 필터링을 통과한 메일은 데이터베이스에 저장되고, SIG는 메일헤더 정보를 메일서버에 보내준다. 헤더정보를 바탕으로 정해진 사용자의 선택에 따라 SIG는 데이터베이스에서 메일원문을 메일서버로 보내거나 삭제한다.




[그림 1] SIG 시스템 네트워크 구성도

### 3.1.2 SIG 구조

SIG는 [그림 2]와 같이 5개의 모듈로 구성되어 있으며 각 모듈별 기능은 다음과 같다.




[그림 2] SIG의 구조

(1) MCM(Mail Capture Module)

내부 네트워크로 들어오는 모든 SMTP패킷을 수집하고 완전한 하나의 메일 형태로 바꾸어 MFM에 전달한다. SMTP가 아닌 패킷은 바로 내부 네트워크로 전달된다.

(2) MFM(Mail Filtering Module)

MCM으로부터 받은 메일에 이미 정의된 정책에 따라 필터링 과정을 수행하여 폐기한다. 필터링에 통과된 메일은 DBM에 전달한다. 필터링 룰은 현재 개개의 시스템이나 사용자를 보호하기 위해 제안되었던 필터링 방식들을 적용한다.

(3) DBM(Data Base Management)

MFM으로부터 메일을 받아 DB에 저장하는 역할이다. 메일을 저장할 때 각 메일을 유일하게 구분하는 Identifier Number를 지정한다. VMS로부터 메일정보 요청이 들어오면 해당메일을 DB에서 찾아서 전해주거나, 삭제한다.

(4) VMS(Virtual Mail Server)

일정시간 마다 DBM에게 새로 저장된 메일의 검색을 요청하고, 검색된 메일은 본 메일서버에 메일헤더 정보만을 보내주고, SMTP 메시지의 DATA 부분[4]에는 Identifier Number를 대체하여 보낸다. SM으로부터 원문 전송 요청이 전달되면 DBM에게 Identifier Number 해당하는 메일을 받아 정상적인 SMTP을 이용하여 메일서버에 원문을 전송한다. 삭제 요청이 전달되면, DBM에게 요청하여 해당 메일을 삭제한다.

(5) SM(SIGP Module)

메일서버로부터 오는 SIGP 패킷을 받아 관련 정보를 SVM에게 전달한다.

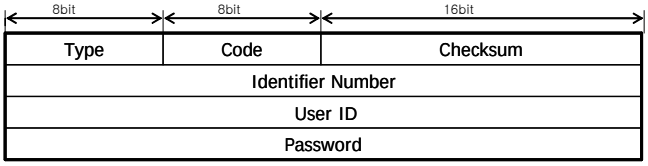
## 3.2 SIGP

### 3.2.1 SIGP 기능

SIGP는 메일의 대한 원문 요청 또는 삭제 메시지를 사용자로부터 메일서버를 통해 SIG에 전달하는 기능을 한다.

### 3.2.2 SIGP 메시지 형식

SIGP 메시지 형식은 [그림 3]과 같이 6개의 필드로 구성된다.




[그림 3] SIGP 메시지

(1) Type 필드

메시지의 종류를 결정한다. 필드 값이 1이면 사용자가 메일서버에 메일원문을 요청하는 메시지이고, 2이면 삭제를 요청하는 메시지이다. 3이면 메일서버가 SIG에 메일원문을 요청하는 메시지이고, 4이면 삭제요청 메시지를 나타낸다.

(2) Code

에러종류를 구분하기 위한 필드이다.

(3) Checksum

에러체크를 위한 필드이다.

(4) Identifier Number

SIG 데이터베이스에 저장되어 있는 메일 중에 사용자가 요청하는 메일을 구분하기 위한 식별자가 들어가는 필드이다.

(5) User ID

인증을 위한 필드로 사용자 ID가 기록된다.

(6) Password

인증을 위한 필드로 사용자 패스워드가 암호화 과정을 거쳐 기록된다.

## 3.3 SIGP 관련 모듈

SIGP 관련 모듈은 SIGP 생성 모듈과 중계 모듈로 구성된다.

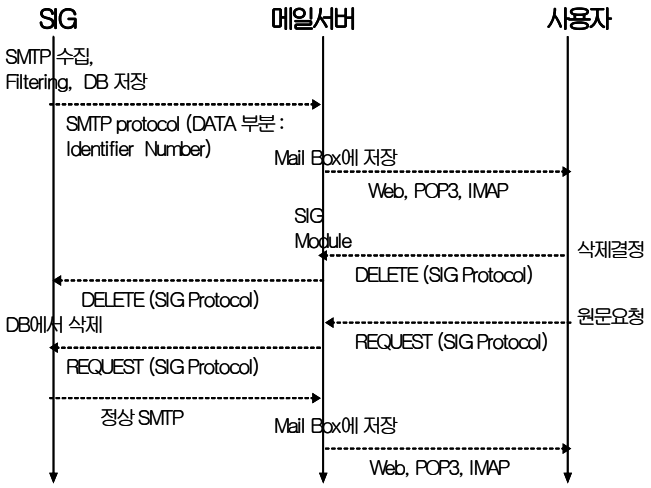
생성 모듈은 사용자의 선택을 SIGP 메시지를 생성하여 메일서버로 전해주는 역할이다. 이 모듈은 사용자의 선택에 따라 원문요청은 Type 필드 값을 1로, 삭제 요청은 2로 설정하고, Identifier Number와 사용자 ID, 그리고 암호화 과정을 거친 패스워드를

기록하여 메일서버로 SIGP를 전송하다.

중계 모듈은 메일서버에 위치하여 사용자로부터 SIGP 메시지를 받고, User ID와 Password 필드를 이용하여 인증과정 거친 후, 인증이 되면, SIGP의 Type 필드를 3 또는 4로 바꾸어 SIG로 전송하는 역할을 한다.

### 3.4 SIG 동작과정

SIG 시스템과 메일서버, 사용자 사이의 전체 동작 과정은 [그림 4]와 같이 이루어진다.




[그림 4] SIG 동작과정

SIG가 필터링을 거쳐 수집된 메일의 헤더정보와 Identifier 정보를 메일서버에 보내면, 사용자는 메일서버에 접속하여 헤더정보를 보고, SIGP을 이용하여 삭제 또는 원문전송 요청을 한다. 메일서버는 사용자의 SIGP 메시지를 받아 인증과정을 거쳐 SIG에 보내주고, SIG는 요청에 따라 삭제 또는 원문 전송을 한다.

## 4. 장점 및 활용

이 시스템의 가장 큰 특징은 2가지로 정리 할 수 있다.

첫 번째는 게이트웨이에서의 필터링이다. 게이트웨이는 내부 네트워크로 들어오는 모든 메일을 수집하여 필터링을 수행하므로 각각의 메일서버로 스팸메일이 도착하지 않는다. 이로 인하여 네트워크의 부하를 방지할 수 있고, 같은 메일을 여러 서버에서 필터링하는 등의 중복된 일을 하지 않아도 됨으로써 각 서버의 자원이 낭비되는 것을 막는다.

두 번째는 사용자의 선택에 의한 메일 전송이다. 게이트웨이에서 메일을 저장한 후 각 메일 서버로 메일헤더 정보만이 전송된다. 그리고 사용자의 선택에 의해서 게이트웨이에서 메일서버로 메일 원문이 전송되거나 삭제한다. 스팸메일의 형태가 다양해지고 있어, 필터링 정책만으로 스팸메일을 완벽하게 차단하는 것은 불가능한 현실에서 사용자의 선택으로 스팸메일 여부를 판단한다면, 필터링 정책에 의한 구분보다는 보다 정확한 구분이 가능할 것이다. 또한 스팸메일은 보통 그림, 문서 파일이 첨부되어 있고 이는 메일 박스의 용량을 많이 차지한다. 이런 문제점은 메일서버의 자원을 낭비하는 것뿐만 아니라, 사용자의 메일 박스 용량을 차지하여 정말 필요한 메일을 받지 못하는 부작용도 생길 수 있다. 이 역시 사용자의 선택되어진 메일만 게이트웨이에서 메일 서버로 전송함으로써 해결되어 질 수 있다.

## 5. 결론 및 향후 계획

스팸메일의 피해를 막기 위한 연구는 최근 몇 년 동안 계속 되었지만, 스팸메일로 인해 낭비되는 네트워크 자원과 메일서버 자원을 보호하고자 하는 연구는 미약했다. 본 논문이 제시한 SIG는 게이트웨이 단에서 SMTP 프로토콜을 이용한 전자 우편 서비스를 감시하여 1차적으로 필터링 및 적절한 정책을 수행하고, 2차적으로 사용자의 선택으로 메일 전송이 이루어짐으로써 내부 네트워크에 있는 메일 서버의 자원과 네트워크의 자원이 스팸메일로 인해 낭비되는 것을 효율적으로 막을 수 있다.

앞으로의 연구에서는 애플리케이션 게이트웨이 특성상 처리속도가 느린 단점을 극복하는 방안을 구상해 볼 것이며, 메일서버에서 이루어지는 SIGP 사용자 인증 부분에 대하여 좀 더 강력한 보안기능 추가에 대하여 연구해 볼 것이다. 또한 게이트웨이를 이용하여 SMTP만이 아닌 내부의 네트워크를 보호하는 통합 보안 시스템으로의 발전도 구상해 볼 것이다.

### 참고문헌

[1] Cristianini N, "An Introduction to Support Vector Machines", Cambridge University, 2000.

[2] 3Com Technical Papers, “Internet Firewalls and Security”, 3COM, July 1996.

[3] Richard S, *TCP/IP Illustrated Volum1*, Addison Wesley, July 1995.

[4] 박영우, *2001년 정보화 역기는 실태 조사 보고서*, 한국 정보보호 진흥원, March, 2001.

[5] 주식회사 데이콤 정규석, *유령 아이디를 이용한 스팸메일 방지 방법*, 대한민국특허청.