Journal of Korea Multimedia Society Vol. 8, No. 1, January 2005(pp. 111-118)

비밀단어의 회신을 이용한 스팸메일 차단 시스템의 구현

고주영[†], 심재창[††], 김현기[†††]

요 약

본 논문에서는 간단하고 빠른 스팸메일 차단 방법으로 등록된 전자우편만 받아 볼 수 있는 스팸메일 차단 시스템을 제안하였다. 사용자는 등록된 전자우편만 수신할 수 있으며 등록되지 않은 전자우편이 수신되면 자동으로 발신자에게 비밀단어를 포함하여 전자우편을 회신하고 발신자가 한번만 비밀단어를 적어 회신하면 등록하는 알고리즘을 구현하였다. 제안한 방법은 정상전자우편 목록만을 관리하므로 간단하여 구현하기 쉽고, 오류율을 최소화 할 수 있으며 DB의 용량이 작은 장점이 있다. 그리고 인트라넷의 경우 전자우편 주소를 비교하기 전에 수신자 도메인 네임 목록을 먼저 검색하여 스팸메일을 빠르게 처리하도록 하였다. 제안된 시스템은 리눅스 시스템에서 procmail, php, IMAP를 이용하여 구현하였으며 실험을 통하여 성능을 확인하였다.

An Implementation of the Spam Mail Prevention System Using Reply Message with Secrete Words

Joo-Young Ko[†], Jae-Chang Shim[††], Hyun-Ki Kim[†††]

ABSTRACT

This paper describes an implementation of the spam mail prevention system using reply message with secrete words. When user receives a new e-mail, the e-mail address is compared with the white e-mail addresses in database by the system. If user receives a new e-mail which does not exist in a white e-mail addresses database, a reply e-mail attached with secrete words is delivered automatically. And the system is compared with the white domains first for intranet environment. It speeds up processing time. Proposed algorithm is required a small database and faster than the black e-mail addresses comparison. This system is implemented using procmail, PHP and IMAP on Linux and the user can manage the databases on the web.

Key words: E-Mail(전자우편), Spam Mail Prevention(스팸메일 차단), White E-Mail(정상 전자우편), Black E-Mail(스팸메일), Secrete Words(비밀단어), White E-Mail DB(정상전자우편 데이터베이스)

1. 서 론

인터넷의 중요한 기능의 하나인 전자우편(E-mail)은 네트워크의 발달과 기능의 편리함으로 우리생활의 중요한 부분으로 자리 잡았다. 그러나 대량의 광고나 불법 음란메일로 인하여 전자우편의 관리에 어려움이 많다. 급속히 증가하는 스팸메일(Spam mail)로 인한 네티즌들이 겪고 있는 피해는 이제 단순한

※ 교신저자(Corresponding Author) : 고주영, 주소 : 경북 안동시 송천동 388번지(760-749), 전화 : 054)820-5911, FAX : 054)823-1630, E-mail : sonice@andong.ac.kr
접수일 : 2004년 4월 16일, 완료일 : 2004년 7월 7일
[†] 준회원, 국립 안동대학교 대학원 정보통신공학과 박사과정
[††] 정회원, 국립 안동대학교 전자정보산업학부 교수 (E-mail : jcshim@andong.ac.kr)
[†††] 정회원, 국립 안동대학교 전자정보산업학부 교수 (E-mail : hkkim@andong.ac.kr)

정보화의 역기능의 차원을 넘어 매우 심각한 사회문제로 인식되고 있다[1].

스팸메일은 다수에게 광범위하게 보내짐으로써 인터넷 서비스제공자 뿐만 아니라 이를 수신하는 개인 사용자와 기업들에 시간적 정신적 경제적 피해뿐 아니라 근로생산성에도 영향을 줄 수 있다[2]. 우리나라는 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제 50조 1항의 누구든지 수신자의 명시적인 수신거부의사에 반하는 영리목적의 광고성 정보를 전송하여서는 아니 된다고 규정하여 Opt-Out 방식에 의한 수신자의 의사에 반하는 스팸메일 송신을 금지하고 있다[3].

스팸메일을 차단하는 기술적인 방법에는 내용기반 스팸메일 차단방법과 전자우편 주소기반 스팸메일 차단 방법으로 나누어 볼 수 있다. 내용기반 스팸메일 차단 방법은 새로 받은 전자우편의 내용을 미리 저장된 전자우편들의 내용과 비교하여 스팸메일을 분류하는 방법이다. 내용기반 방식의 스팸메일 필터를 사용할 경우 통계에 사용된 전자우편과 유형이 다른 정상메일을 스팸메일로 분류하는 오류율(False-positive)이 발생한다. 그러므로 스팸메일로 분류된 전자우편들을 항상 확인해야 하는 번거로움이 있다. 그리고 데이터베이스에 많은 전자우편 자료들이 있을수록 좋은 결과를 나타내므로 사용자는 항상 많은 전자우편 자료를 저장해 두어야 한다.

전자우편의 주소를 기반으로 스팸메일을 차단하는 방법은 스팸메일 주소(Black e-mail address)를 차단하는 방법과 정상메일 주소(White e-mail address)만 받아 보는 방법이 있다. 스팸메일 주소를 차단 할 경우 스팸메일 주소가 자주 변경되고 일회성이 많아 스팸메일 주소의 양이 계속 증가되어 데이터베이스가 커지는 단점이 있다[4]. 정상메일 주소만 받아 보는 방법은 스팸메일을 수신하지 않는 가장 확실한 방법이지만 데이터베이스에 등록되지 않은 발신자의 전자우편을 받을 수 없고 새로운 사용자를 계속 수동으로 입력해야 하는 문제점이 있다.

본 논문에서는 옵트인 방식에 기반을 둔 비밀단어의 회신을 이용한 스팸메일 차단 시스템을 구현하였다. 제안한 방법은 본인의 서버에 등록된 전자우편만 수신할 수 있으며 등록되지 않은 전자우편이 수신되면 자동으로 발신자에게 비밀단어를 포함한 전자우편을 회신을 한다. 전자우편 수신을 위한 데이터베이스에 등록하기 위해 단 한번 발신자가 제목에 비밀단어를 적어 보내고, 이를 회신하도록 요청해서 확인하는 알고리즘을 제안하였다. 제안한 방법은 정상메일만 받아 볼 수 있어 스팸메일을 따로 관리하지 않아도 되며 정상메일 주소 데이터베이스만을 관리하므로 데이터베이스 용량이 줄고 문자 비교에서 속도가 향상된다. 또한 웹에서 사용자가 직접 정상메일 주소와 스팸메일 주소를 관리 할 수 있도록 하였으며 도메인 네임을 관리 할 수 있도록 구현하였다. 그리고 인트라넷에서 전자우편 주소를 비교하기 전에 수신자 도메인 네임 목록을 먼저 검색하여 스팸메일을 처리하도록 하였다.

2. 스팸메일과 차단방법

2.1 내용기반 스팸메일 차단방법

스팸메일은 원래 소비자 정보를 수신자들에게 전달하는 순기능을 가지고 있다. 그러나 스패머(Spammer)들에 의해 대량으로 광범위하게 보내짐으로 네티즌에게 개인의 경제적 심리적 부담과 개인정보 침해의 피해를 준다. 스팸메일의 유형이 단순 광고용인 경우보다 음란 정보 스팸메일이 급격이 증가하여 사용자들에게 불쾌감을 주고 단순히 시간낭비, 정신적 피해뿐만 아니라 직장인의 경우 근로생산성 문제에까지 영향을 줄 수 있는 것으로 볼 수 있다.

내용기반에 의한 스팸메일 차단방법은 미리 분류되어 저장된 전자우편들의 정보를 바탕으로 새로 받은 전자우편을 분류하는 방법이다. 규칙기반의 Cohen의 RIPPER 알고리즘[5]과 Mooney가 설명한 나이브 베이지안 알고리즘[6]을 이용한 필터는 많은 연구가 되고 있다.

규칙기반 알고리즘은 키워드 맞춤 규칙에 의해 설계되며 전자우편 분류와 필터링을 위해 적합하며 전자우편을 읽는 프로그램에 이미 이런 종류의 분류 규칙을 사용하고 있다. 그러나 메일 사용자에 따른 규칙의 배열을 변환하는 문제가 있다. 국내에서도 베이지안과 메시지 규칙을 이용한 스팸메일 필터링에 관한 연구[7]와 통계적 학습이론에 기반을 둔 방법으로 스팸메일을 차단하는 연구[8]가 진행되었다.

현재 많은 사람들이 사용하고 있는 베이지안 알고리즘을 사용한 POPfile[9]은 매우 효과적인 차단 필터로 알려져 있다. 표 1은 POPfile을 4개월간 실제 사용한 예이다.

표 1에서와 같이 추론 알고리즘을 이용한 스팸필터는 미리 저장되어 분류된 전자우편의 정보를 이용하여 새로 받은 전자우편을 분류한다. 미리 저장되어진 전자우편이 많을수록 스팸메일을 분류하는 정확도가 높아진다. 그러므로 항상 많은 메일을 보관해야 하고 추론과정을 이용하므로 전자우편을 잘못 분류하는 오류가 발생하여 사용자는 항상 스팸메일 함을 다시 확인하여 전자우편을 다시 분류해주는 과정을 거쳐야 한다. 국내 인터넷 서비스 업체의 스팸메일 차단방법을 살펴보면 표 2와 같다.[10]

대부분의 국내 인터넷 서비스 업체에서 전자우편 분류규칙과 전자우편 주소 차단 및 허용 등의 서비스를 제공하고 있다. 그러나 전자우편 분류 규칙이나 주소를 차단 또는 허용할 수 있는 정보의 개수가 제한되어 있는 경우가 많다. 또한 수신 거부할 주소를 서비스 제공업체에서 온라인으로 업데이트 하는 방법이 사용되고 있다[11]. 이 방법은 음란 메일을 받지 않을 수 있는 좋은 방법이지만 처리 시간이 지연되고 중요한 전자우편이 스팸메일로 분류되었을 경우 복구하는데 어려움이 발생한다.

표 1. POPfile을 사용한 예

	1개월	2개월	3개월	4개월
E-mail	1,081	1,794	2,369	4,471
False Positive	124	166	185	254
Accuracy	88.5%	90.7%	92.2%	94.3%

표 2. 국내 인터넷 서비스 업체의 스팸메일 차단방법

서비스 업체	스팸 등급설정	전자우편분류규칙	주소 수신거부 및 허용	기타
chol.com	○	○	○	도메인 수신거부 및 허용
empas.com	○	○	○	실명 인증
hanafos.com		○	○	
hanmail.net		○	○	온라인 우표제
hotmail.com			○	
kebi.com		○	○	
korea.com	○	○	○	
outlook	○	○	○	
yahoo.co.kr	○	○	○	벌크메일함

2.2 전자우편 주소기반 스팸메일 차단방법

전자우편 주소기반으로 스팸메일을 차단하는 기본적인 방법으로 스팸메일 주소를 차단하는 방법과 정상메일 주소만 수신하는 방법으로 시스템을 구성할 수 있다. 스팸메일 주소를 차단하는 방법은 현재 가장 많이 사용되고 있는 방법이다. 그러나 스패머들이 발신자의 주소를 일회성으로 사용하는 경우가 많아 스팸메일 주소 데이터베이스가 계속 증가하게 된다. 그러므로 스팸메일 차단보다는 정상메일 주소 수신 방법이 데이터베이스 활용 면과 처리 속도측면에서 효과적이다.

그림 1은 정상메일을 통과시키는 방법이다. 이 방법은 자신이 원하는 전자우편만 받을 수 있는 가장 좋은 방법이지만 전자우편 주소가 등록되지 않은 경우 전자우편을 받을 수 없기 때문에 새로운 사용자를 계속 직접 입력해야 하는 불편함이 있다. 일반적으로 수신된 전자우편에서 정상메일 주소를 추출하여 전자우편 주소 데이터베이스와 비교한다.

본 논문에서는 스팸메일 차단 방법으로 사용자가 원하는 전자우편만 받아 볼 수 있게 정상메일만 통과시키는 방법을 사용하였다. 정상메일만 허용할 경우 등록되지 않은 전자우편은 수신할 수 없게 되는 것을 방지하기 위해 등록되지 않은 전자우편이 수신되면 서버에서 자동으로 비밀단어를 포함하여 메시지를 발신자에게 회신하도록 하였다. 발신자가 비밀단어를 제목 란에 입력하여 전자우편을 보내면 그 전자우편 주소는 자동으로 정상메일 데이터베이스에 저장되도록 하였다. 전자우편 수신자가 한번 확인해야 하는 번거로움이 있으나 정상메일 주소 데이터베이스

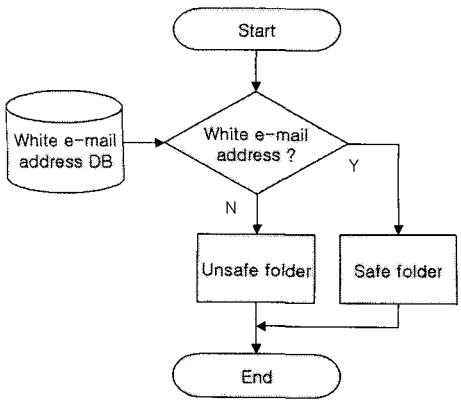


그림 1. 정상메일 주소 통과 시스템

에 등록되지 않은 주소만 회신하여 불편함을 최소화 하였다. 스팸메일 방지를 위해 발신자 확인 방법에 버튼을 누르는 방법과 이미지를 이용하는 방법 등을 사용할 수 있으나 정형화된 방법을 사용하였을 때 스패머들이 이를 역이용할 가능성이 있어 사용자가 직접 비밀단어를 입력하고 변경하여 메시지를 적음으로써 스패머들이 이를 역이용할 수 없게 하기 위해 비밀단어를 입력하는 방법을 사용하였다.

3. 비밀단어의 회신을 이용한 스팸메일 차단 시스템

3.1 시스템의 구성

전자우편 주소는 userID@domain-name.com 로 구성된다. 제안하는 방법은 수신하기 원하는 전자우편 주소가 저장된 정상메일주소 데이터베이스를 만들고 수신된 전자우편에서 주소를 추출하여 데이터베이스의 전자우편주소와 비교하는 방법이다. 정상메일주소를 비교하는 방법은 스팸메일을 수신하지 않는 좋은 방법이지만 데이터베이스에 등록되지 않은 처음 보내는 발신자의 전자우편을 받을 수 없는 불편을 줄이기 위해 비밀단어의 회신을 이용하여 새로운 전자우편 주소를 자동으로 등록할 수 있도록 하였다. 그림 2는 수신 이메일 주소와 비교하여 비밀단어 회신을 이용한 프로그램 수행과정이다.

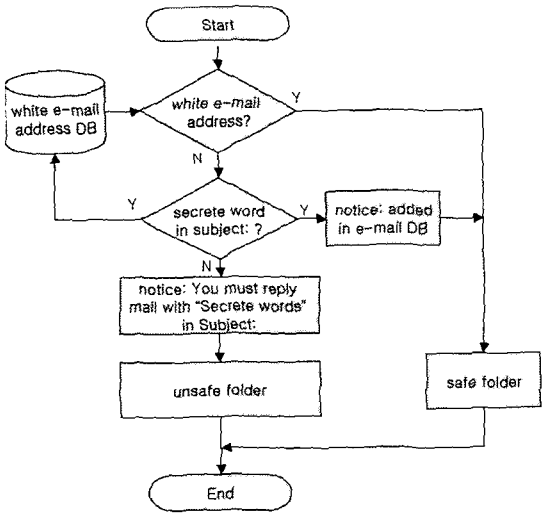


그림 2. 수신 이메일 주소와 비교하여 비밀단어 회신을 이용한 프로그램 수행과정

또한 인트라넷 환경에서 전자우편을 사용하는 대학이나 기업에서 적용할 때 주로 사용하는 도메인(white domain name)을 수개 설정하고 미리 검사하면 처리속도가 더 빠르다. 특히 리포트를 받거나 다량의 메일을 인트라넷에서 활용하는 경우 이 방법이 유용하다. 그림 3은 특정 도메인을 먼저 비교하는 프로그램의 전체수행과정을 나타낸다.

제안된 시스템을 리눅스 시스템에서 구현하였다. Sendmail과 Procmail 및 IMAP[12] 데몬이 실행되는 환경에서 작동된다. Sendmail은 리눅스 시스템에서 전자우편 서버프로그램으로 가장 많이 사용하는 프로그램으로 사용자로부터 전자우편 송신 요청을 받아 외부 네트워크로 전송해주는 메일서버이다.

Procmail[13]은 외부에서 Sendmail을 통해 들어오는 메일을 필터링 할 때 주로 사용된다. Procmail을 사용하면 메일의 헤더정보를 쉽게 바꿀 수 있고 본문의 내용을 각각의 문자 셋에 따라 코드변환 할 수 있다. Procmail은 그 자체로서 실행되지 않기 때문에 시스템 관리자가 sendmail.cf에 포함시키거나 .forward파일을 두어 실행한다. *procmail*을 포함한 메일 송수신 과정은 그림 4와 같다.

스팸메일 차단 시스템의 처리 과정은 전자우편이 도착하면 먼저 발신자의 도메인이 정상메일 데이터베이스에 있는 경우 Safe folder로 보내고, 그렇지 않은 경우 다음 단계인 사용자 확인을 한다. 전자우편

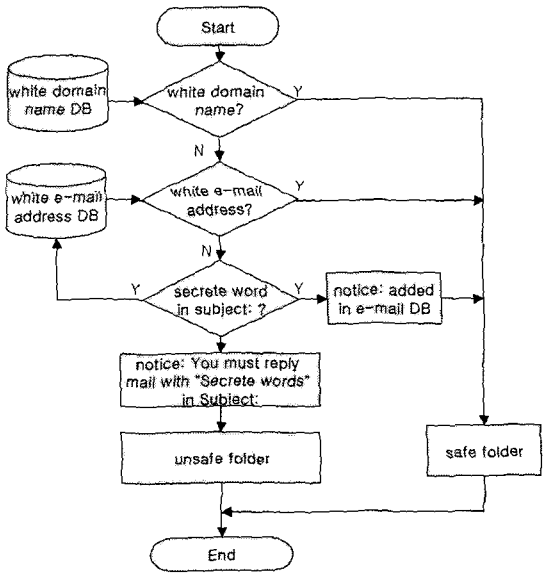


그림 3. 시스템의 구성도

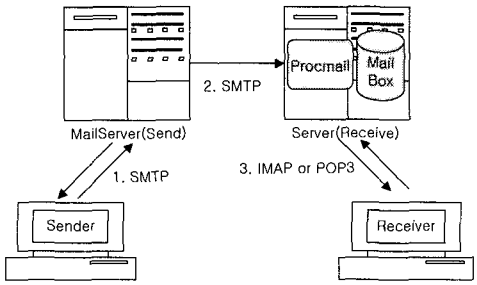


그림 4. Procmail을 포함한 메일 송수신 과정

주소를 추출하여 정상메일 주소 데이터베이스에 등록되었는지 확인한다. 등록된 경우 스팸메일 주소 데이터베이스를 확인해서 스팸메일 주소가 아닌 경우 Safe folder로 보내고 그렇지 않으면 스팸이므로 버린다. 정상메일 주소 데이터베이스에 등록되어 있지 않는 경우 비밀단어를 확인한 다음 비밀단어가 있으면 Safe folder로 보내고 비밀단어가 없으면 회신 비밀단어(/.returnsecretewords.txt)를 붙여서 Sendmail을 이용하여 전자우편을 반송하고 Unsafe folder로 보낸다.

만약 발신자가 반송메시지를 받고 보내온 비밀단어를 제목에 입력한 후 다시 전자우편을 보내면 스팸메일 차단 시스템에서 자동으로 정상메일 주소(White e-mail address DB)가 데이터베이스에 등록되고 알려준다. 그림 5는 전자우편이 수신된 후 처리되는 과정을 나타낸 시스템의 구성도이다.

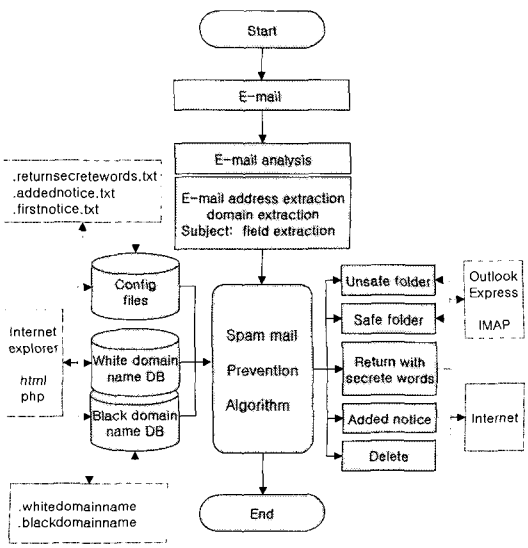


그림 5. 전자우편 주소기반의 스팸메일 차단 시스템의 구성도

3.2 설정 파일의 변경을 위한 사용자 인터페이스

제안된 시스템은 비밀단어(.returnsecretewords.txt), 수신 이메일 주소 등록 알림 메시지(.added_notice.txt) 그리고 이메일 주소 등록을 위한 첫 번째 알림 메시지(.first_notice.txt)를 사용자가 직접 관리할 수 있도록 서버의 사용자 폴더에서 PHP 프로그램으로 구현하였다. 또한 수신 도메인 네임 목록과 수신 이메일 주소 목록을 직접 관리 할 수 있도록 사용자 인터페이스를 구현하였다. 그림 6은 파일 내용 설정 사용자 인터페이스의 구성도이다.

제안된 시스템을 IMAP(Internet Message Access Protocol)[12]와 연동함으로 서버에 있는 메시지를 자신의 로컬 컴퓨터에 있는 것처럼 사용하기 쉽게 된다. IMAP는 클라이언트가 메일 서버에서 메일을 읽기 위한 인터넷 표준 프로토콜이고 기본적인 역할은 POP3과 동일하다. 그림 7은 IMAP에서 서버와 클라이언트 구성을 나타낸다.

그림 8은 IMAP를 활용한 메일서버와 Outlook Express의 구성도이다. POP3보다 유연성이 뛰어난 IMAP를 이용할 경우 서버의 unsafe folder와 safe folder를 Outlook Express에서의 폴더처럼 사용할 수 있고, 다른 컴퓨터에서 접속할때도 웹 메일처럼 활용이 가능하다.

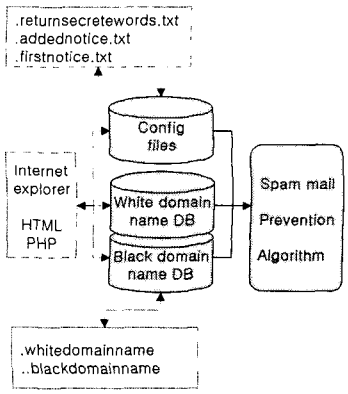


그림 6. 파일 내용 설정 사용자 인터페이스

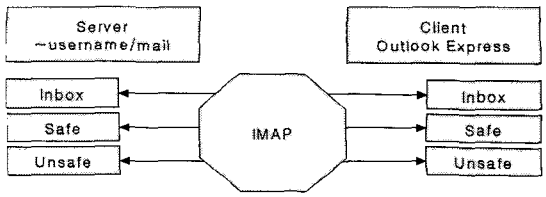


그림 7. IMAP에서 서버와 클라이언트의 구성

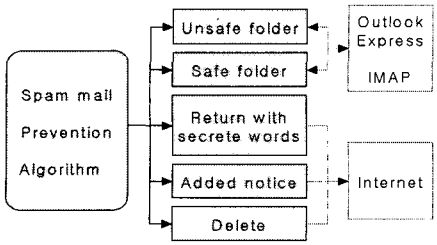


그림 8. IMAP를 활용한 메일서버와 Outlook Express

4. 실험 및 고찰

시스템의 실험은 구현된 프로그램을 실험용 서버 컴퓨터에 설치하고 설계된 내용과 동일하게 작동하는가를 확인하였다. 서버는 리눅스 시스템이 설치되어 있으며 한컴리눅스 3.1이 설치되어 있다. 4개월 동안 사용해 본 결과 제안된 방법을 적용한 결과 알고리즘이 잘 작동 되었고 실험 기간 동안 평균 90통의 메일이 수신되었고, 조사결과 약 90%가 스팸메일이며 비밀단어가 된 회송 메일을 반송을 하지 않아 등록되지 않고 폐기가 되었다. 수동으로 등록된 240명을 제외하고 이 기간 동안 새로 등록한 경우 95명이 새로 등록되었다. 처리속도는 이메일의 처리에 시간이 거의 소요되지 않아 문자열의 길이를 통하여 비교하였다.

제안된 방법을 시뮬레이션 하기 위해서는 리눅스 시스템에서 Sendmail, Procmail, IMAP 데몬이 실행되어야 한다. 그리고 본인의 홈 디렉토리에 .procmailrc, .forward, .add_notice.txt, .first_notice.txt, .returnsecreteword.txt를 복사하고 log 파일을 저장하기 위한 .procmail 폴더를 생성한다. 또한 아웃룩 익스프레스에서 IMAP를 설정하고 IMAP에서 Safe, Unsafe 폴더 작성한다. 프로그램의 설치가 되었으면 사용자 설정 인터페이스를 작성한다.

새로운 발신자의 메일이 처음 도착하였을 때 전자우편 주소를 자동으로 등록하기 위해서 회신비밀단어를 포함한 반송메시지를 보내어 발신자를 확인하여 자동으로 등록되도록 하였다. 사용자 인터페이스 설정 입력창에서 회신 비밀단어, 알림 글 그리고 전자우편 주소가 등록되었음을 알리는 메시지를 입력할 수 있다. 발신자가 전자우편을 보냈을 경우 비밀단어는 사용자가 한 개 또는 여러 개 지정할 수 있고 여러 개로 지정을 할 때 단어 사이를 '|'로 구분하도록 하였다. 비밀단어가 두 개 이상일 때 각각의 비밀단어로 수신된 전자우편은 비밀단어에 따라 각각의 폴더로 구분되어 분류할 수 있다. 그림 9는 사용자 인터페이스 설정 입력창이다. 그림 9에서 Secrete words란에 비밀단어를 입력하고 Added notice란에 발신자의 사용자 계정이 등록되었음을 알리는 메시지를 입력한다. 그리고 First notice란에 비밀단어와 메시지를 보내는데 이때 사용자가 임의로 비밀단어와 메시지를 입력할 수 있기 때문에 스패머들의 기계에 의한 비밀단어의 추출이 어렵다.

발신자가 메시지를 받고 회신비밀단어를 입력한 다음 다시 전자우편을 보내면 스팸메일 차단시스템이 비밀단어가 맞는 경우 정상메일 주소 목록에 자동 저장된다. 그리고 발신자에게 전자우편 주소가 등록되었음을 알린다.

그림 10은 발신자가 반송 메시지를 받은 결과이다.

그림 11은 사용자가 등록되었음을 알리는 메시지이다. 발신자가 메시지를 받고 반송비밀단어를 입력한 다음 다시 메일을 보내면 스팸메일 차단시스템이 비밀단어를 확인하고 비밀단어가 맞는 경우 정상메일 주소 목록에 자동 저장된다.

정상메일 주소 목록에 등록된 발신자가 전자우편을 보냈을 경우 반송 과정 없이 바로 전자우편이 Safe folder로 전달되고 발신자에게 전자우편이 전달되었음을 알린다. 제안된 방법은 자신이 원하는 전자

그림 9. 사용자 설정 인터페이스 입력 창

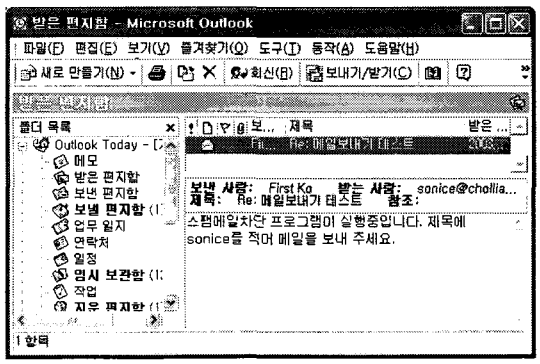


그림 10. 발신자가 반송메시지를 받은 결과

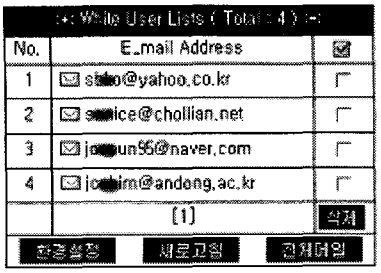


그림 11. 등록된 정상메일 주소 목록

우편만 받을 수 있고 주소가 등록되지 않은 전자우편이 도착했을 때에는 시스템이 자동으로 비밀단어를 포함한 반송메일을 보내어 확인함으로 자동으로 주소를 등록할 수 있으며 스팸메일로 잘못 분류하는 오류율을 최소화 할 수 있다. 또한 정상메일 주소 데이터베이스에 등록되지 않은 주소인 경우만 발신자 확인을 하여 불편함을 최소화하였다.

제안된 스팸차단 방법과 기존의 스팸 차단 방법을 비교 평가하였다. 수신자에게 100통의 전자우편이도착하였다고 가정한다. 이때 전자우편 userID@domain.com에서 평균 문자수는 userID 문자 5.8자와 domain 문자 15.5자를 합하여 21.3자 이다. 이는 palgong.knu.ac.kr과 andong.ac.kr 도메인의 통계이다. 예를 들어 수신 전자우편 주소가 1000개이고 Black e-mail address가 5000개이면 비교되는 문자수는 표 3과 같다.

또한 정상도메인 네임을 먼저 비교하고 정상메일 주소를 비교하면 인트라넷에서 속도가 빨라진다. 100통의 전자우편이 도착했을 때 2개의 정상도메인 네임을 먼저 비교하여 20%의 전자우편 주소가 제외되었다면 100통의 전자우편 주소 중 20%를 제외한 전자우편 주소(80통)를 비교하면 표 4와 같은 결과를 얻을 수 있다. 인트라넷에서는 리포트 접수 등이 있는 경우는 전자우편 량이 증가된다.

이와 같이 제안된 방법은 정상메일 주소 데이터베이스를 등록함으로 데이터베이스의 용량을 줄이고 등록되지 않은 사용자는 비밀단어 회신을 통하여 자동으로 등록시키고 스팸메일 차단 처리 속도를 증가시킬 수 있다.

5. 결 론

본 논문에서 제안한 스팸메일 차단 방법은 전자우편이 도착하면 등록된 정상메일 주소를 비교하여 차단하고 등록되지 않은 전자우편이 도착하면 자동으로 비밀단어를 포함한 반송메시지를 보내어 발신자를 확인하는 과정을 거침으로 사용자가 일일이 수신가능 메일을 등록해야하는 불편함을 줄였다. 스패머들이 자동 전자우편 보내기 프로그램을 사용하여 발신한 스팸메일은 반송비밀 단어를 일일이 확인해 줄 수 없으므로 스패머가 보내는 전자우편은 받지 않을 수 있다. 그리고 확인을 거치지 않은 스팸메일은 받은 전자우편함에 전달되지 않기 때문에 기존의 필터를 사용하는 경우 일일이 스팸메일함을 다시 확인해야 하는 번거로움을 없앨 수 있다.

또한 인트라넷 환경의 학교나 기업체에 적용할 때 정상도메인 네임을 미리 검사한 후 정상메일 주소와 비교하는 방법으로 스팸메일 차단 처리속도를 증가시켰다. 또한 사용자가 웹 브라우저에서 설정 파일 사용자 인터페이스를 관리할 수 있으며 IMAP와 연동해서 Outlook Express에서 사용할 수 있도록 하였다.

중요한 전자우편을 반송 없이 받기 위해서 미리 직접 등록할 수 있도록 사용자 설정 인터페이스를 구현하여 발신자가 확인을 하는 과정을 최소화하였다. 제안된 방법은 개인의 스팸메일 차단 뿐 아니라 대량의 전자우편이 통과되는 전자우편 서버 시스템

표 3. Black e-mail과 White e-mail에서 비교 문자 수

Methods	Comparison characters	Totals
Black e-mail comparison	100통×5000개×21.3자	10,650,000
White e-mail comparison	100통×1000개×21.3자	2,130,000

표 4. White domain을 먼저 필터링하고 White e-mail을 비교 검사하는 경우

비교대상	문자 비교 수	계
White domain비교	100통×15.5자×2개	1,707,100
White e-mail비교	80통×21.3자×1000개	

에 적용하기에 적합하다. 앞으로 대용량 전자우편 수신 서버에 적용할 수 있도록 시스템을 구현하고 안정성 있는 인증기술등을 통합하여 다양한 환경에서 스팸메일 차단을 할 수 있도록 계속적인 연구가 필요하다.

참 고 문 헌

[1] 2003년 정보화 역기능 실태조사서, 한국정보보호진흥원, 2003.

[2] 주덕규, 스팸메일의 현황 및 대책, 정보통신윤리, pp. 8-17, 2003.

[3] 정보통신망이용촉진및정보보호등에 관한 법률, 대한민국 법률 제6360호, 2004.

[4] H. Drucker, D. Wu. and V. N. Vapnik, "Support Vector Machines for Spam Categorization," *IEEE Transactions on Neural networks,* Vol. 10, No. 5, pp. 1048-1054, 1999.

[5] W. W. Cohen, "Learning Rules that Classify E-Mail," *AAAI Spring Symposium on Machine Learning in Information Access,* pp. 18-25, 1996.

[6] Jefferson Provost, "Naive-Bayes vs. Rule-Learning in Classification of Email," *The University of Texas at Austin, Artificial Intelligence Lab. Technical Report* AI-TR-99-284, 1999.

[7] 조한철, 조근식, "나이브 베이지안 분류자와 메세지 규칙을 이용한 스팸메일 필터링 시스템", 한국정보과학회, 제29회 춘계학술대회, pp. 223-225, 2002.

[8] 민도식, 송무희, 손기준, 이상조, "SVM 분류 알고리즘을 이용한 스팸메일 필터링," 한국정보과학회 03 봄 학술발표논문집(B), pp. 552-554, 1598-5164, 2003.

[9] "POPFile", http://popfile.sourceforge.net

[10] "불법스팸대응센터", http://www.spamcop.or.kr/indexb_5.html

[11] "CleanSpam", http://www.cleanspam.co.kr/event/kisa/index.html

[12] M. Crispin, "Internet Message Access Protocol-Version4," http://www.imap.org/papers/docs/rfc 3501.html

[13] "Procmail version 3.22 released," http://www.procmail.org

고 주 영

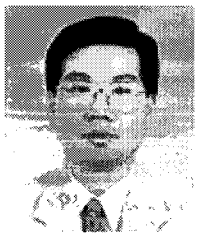
1994년 효성여자대학교 의류학과(석사)
2002년 안동대학교 멀티미디어공학과(석사)
2004년~현재 안동대학교 정보통신공학과 박사과정

관심분야 : 멀티미디어응용, 원격교육

심 재 창

1993년 경북대학교 전자공학(박사)
1997년~1999년 미국 IBM Watson 연구소 연구원
1994년~현재 국립 안동대학교 전자정보산업학부 교수

관심분야 : 영상처리, 패턴인식, 컴퓨터비젼



김 현 기

1986년 경북대학교 전자공학과(공학사)
1988년 경북대학교 대학원 전자공학과(공학석사)
2000년 경북대학교 대학원 전자공학과(공학박사)
1988년~1995년 한국전자통신연구원 선임연구원
1995년~2001년 경남정보대학 전자정보학부 조교수
2002년~현재 국립 안동대학교 전자정보산업학부 교수
관심분야 : 멀티미디어 시스템, 원격교육, 멀티미디어응용