DOI: 10.3745/KIPSTC.2011.18C.2.061

수집과 빈도 분석을 이용한 인터넷 게시판의 스팸 메시지 차단 방법

김 태 희[†] · 강 문 설[††]

요 약

인터넷 환경의 빠른 발전과 함께 널리 사용되고 있는 인터넷 게시판이 기본적인 의사소통 수단으로 정착되고 있으나, 불특정 다수로부터 게시되는 스팸 메시지의 증가로 피해 규모가 날로 증가하고 있다. 현재 스팸 메일을 차단하기 위한 다양한 차단 방법들이 제안되고 있으나 게시판에 자동으로 등록되고 있는 스팸 메시지를 차단하는 방법에 대한 연구는 미미한 실정이다.

본 논문은 인터넷 게시판에 등록되는 스팸 메시지를 수집하여 메시지의 특성과 빈도를 분석하고 차단 규칙을 생성하여 차단하는 단계로 구성된 게시판 스팸 메시지 차단 방법을 제안하였다. 인터넷 게시판의 데이터베이스에 저장되는 모든 메시지를 대상으로 스팸 메시지를 수집하고, 수집한 스팸 메시지를 분석하여 스팸 메시지를 정의할 수 있는 정규화된 규칙을 생성한 후, 이 규칙을 이용하여 등록된 메시지에 대해 스팸 메시지를 검사하여 차단한다. 제안한 방법은 수집되는 스팸 메시지의 정보를 이용하여 다양한 유형의 스팸 메시지를 차단할 수 있으며, 변화하는 스팸 메시지의 형태에 대해서도 유연하게 대처할 수 있는 구조를 가지고 있다.

키워드 : 인터넷 게시판, 스팸 메시지, 수집, 빈도 분석, 차단

Spam Message Filtering for Internet Communities using Collection and Frequency Analysis

Kim, Tae-Hee[†] · Kang, Moon-Seol[††]

ABSTRACT

Even though internet community is becoming the basic communication tool with rapidly changing internet environment, its damage is on the rise due to increasing spam messages of unspecified individuals. Currently, various methods to block spam mails, but studies on block spam messages from being automatically posted in community are still insufficient.

This study suggested methods to block spam messages in internet community by collecting spam messages posted in internet community to analyze characteristics and frequencies of the messages and create block regulations. It collects spam messages of all messages saved in database of internet community, analyze the collected messages to create normalized rules that can define spam messages, and inspect spam messages among posted messages by using the regulations to block them. The suggested method has a structure that can block various types of spam messages by using information of spam messages collected and flexibly deal with changing spam message types.

Keywords : Internet Community, Spam Message, Collection, Frequency Analysis, Block

1. 서 론

유무선 인터넷 환경의 급속한 발달은 커뮤니티 패러다임을 변화시키고 있으며, 원격 이용자들 간의 의사소통, 광고, 홍보 및 마케팅 등을 담당하는 커뮤니티 서비스들은 인터넷의 필수 요소로 빠르게 정착되고 있다. 이러한 서비스들의 대부분은 원활한 의사소통 환경을 제공하고, 대량의 정보 및 지식의 전파와 공유를 가능하게 하는 등 인터넷 순기능의 활성화에 커다란 기여를 하고 있다. 그러나 스팸 메시지 또는 스팸 광고라 불리는 메시지들은 상업적 광고, 청소년 유해물, 비방과 욕설, 명예훼손과 모욕 등을 주제로 수신자의 의사와 상관없이 대량으로 송신되거나 등록되고 있다. 따라서 의사소통의 저해, 자원 및 인력의 낭비, 그리고 인터

※ 이 연구는 2010년도 광주대학교 대학 연구비의 지원을 받아 수행되었음.
† 종신회원 : 동신대학교 디지털콘텐츠학과 부교수
†† 종신회원 : 광주대학교 컴퓨터공학과 교수(교신저자)
논문접수 : 2010년 9월 8일
수 정 일 : 1차 2010년 11월 15일
심사완료 : 2010년 12월 6일

넷 윤리 문제를 야기하는 스팸 메시지들은 인터넷 커뮤니티 환경에서 반드시 해결되어야 할 과제로 부각되고 있다[1, 2].

의사소통 수단이 이메일로 대표되던 초기의 인터넷 환경에서 스팸 메시지는 이메일 서비스에 국한되는 문제였다. 그러나 인터넷 게시판, 인스턴스 메시지 및 휴대폰 메시지 등 의사소통을 위한 다양한 인터넷 서비스가 제공됨에 따라 스팸은 이러한 서비스들을 매개체로 그 양과 피해를 급증시키고 있는 대표적인 인터넷 역기능 중의 하나이다. 현재 스팸 메시지는 송신되는 이메일의 40% 이상을 차지하며, 이로 인해 매해 약 230억 원 이상의 피해가 발생하고 있다. 또한, 스팸은 매해 인스턴스 메시지의 형태(SPIM)로 30억 개 이상의 스팸 메시지가 송신되고 있으며, 이외의 다양한 형태로 인터넷 이용자에게 송신되는 등 인터넷 전반에 걸쳐 막대한 피해를 발생시키고 있다[1, 3].

게시판 스팸(Community Spam)은 인터넷 게시판, 블로그 등과 같은 인터넷 게시판에 게재되는 스팸 메시지로, 피해가 급격히 증가하고 있는 주목할 필요성이 있는 스팸 중 하나이다. 게시판 스팸은 일반적인 스팸에 비해 적은 비용으로 불특정 다수의 인터넷 이용자에게 의사소통의 저해, 인터넷 게시판의 신뢰도 저하 등의 문제를 일으키는 경향이 있다. 일반적으로 이메일 스팸은 이메일과 인스턴스 메시지 등의 형태로 단일 스팸 메시지가 단일 인터넷 이용자에게만 피해를 가하는 형태이다. 그러나 게시판 스팸은 동일한 비용의 단일 스팸 메시지가 해당 인터넷 게시판에 접근하는 다수의 인터넷 이용자와 인터넷 게시판 관리자에게 피해를 입히고 있다. 이와 함께 게시판 스팸은 인터넷 게시판의 접근이 용이해지고, 인터넷 이용자의 사회 참여와 의견 게재가 활발해지는 인터넷 환경에서 그 양과 피해가 급격하게 증가하는 추세에 있다.

인터넷 환경 전반에 걸쳐 스팸 메시지로 인하여 발생하는 문제가 급격하게 증가하고 있지만, 스팸 메시지 차단을 위한 연구는 아직까지 이메일 서비스에 집중되어 있다. 높은 정확도와 낮은 오탐지율을 목표로 이메일 서비스 상의 스팸 메일 차단 방법들에 대한 연구들이 국내외에서 활발하게 진행되고 있다[4,5,6,7,8,9]. 그리고 인스턴스 메시지 서비스와 같이 최근에 널리 보급된 서비스들에 대한 스팸 차단에 관한 연구[10]가 현재 꾸준하게 증가하고 있으나 아직은 미비한 수준이며, 인터넷 게시판의 스팸 차단에 관한 연구는 스팸 메일 서비스나 인스턴스 메시지 서비스의 스팸 연구에 비해 매우 미미한 실정이다[1].

본 논문은 게시판에 스팸 메시지를 남기는 기술이 갈수록 정교해지고 있기 때문에 스팸 게시물을 100% 구분하여 차단할 수 있는 방법은 존재하지 않는다고 가정하고, 가능한 방법을 적용하여 새로운 유형의 스팸 메시지에 대처하면서 꾸준하게 관리하면 대부분의 스팸 메시지를 차단할 수 있도록 지원하는 인터넷 게시판의 스팸 메시지 차단 시스템을 설계하여 구현하였다.

인터넷 게시판의 데이터베이스에 저장되는 모든 메시지를 대상으로 스팸 메시지를 수집하고, 수집한 스팸 메시지의 특성과 빈도를 분석하여 스팸 메시지를 정의할 수 있는 정규화된 스팸 메시지 차단 규칙을 생성한 후, 이 규칙을 이용하여 등록되는 메시지에 대해 스팸 메시지를 검사하여 차단한다. 대학 및 학과의 홈 페이지 게시판에 등록되는 메시지를 대상으로 정상 메시지와 스팸 메시지로 분류하고, 스팸 메시지는 메시지의 특성에 따라 5 가지의 분야로 세분화하며, 이들 스팸 메시지로부터 스팸 정보 데이터베이스와 스팸 단어 데이터베이스를 생성하여 스팸 메시지 차단에 사용한다. 제안한 스팸 메시지 차단 방법은 수집되는 스팸 메시지의 정보를 이용하여 다양한 유형의 스팸 메시지를 차단할 수 있으며, 변화하는 스팸 메시지의 형태에 대해서도 유연하게 대처할 수 있는 구조를 가지고 있다.

본 논문의 구성은 다음과 같다. 2장에서 스팸 메시지의 등록 및 차단 방법에 대한 관련 연구를 살펴보고, 3장에서 스팸 메시지 수집, 메시지 특성과 빈도 분석, 차단 규칙 생성 및 차단하는 단계로 구성되는 스팸 메시지 차단 방법을 설명한다. 4장에서 스팸 메시지 차단 시스템의 구현 및 성능 평가 결과를 기술하며, 5장에서 결론 및 향후 연구 방향을 기술한다.

2. 관련 연구

2.1 스팸 메시지 등록 방법

인터넷 게시판의 스팸 메시지는 시간과 장소를 가리지 않고 등록되고 있다. 어떤 날은 수십 개의 스팸 메시지가 올라오기도 하고, 어떤 때에는 마우스를 클릭하자마자 스팸 메시지를 게시한 홈 페이지로 바로 이동해버려 지울 수도 없는 스팸 메시지가 등록되기도 한다. 이러한 스팸 메시지가 적게는 수천 개에서 많게는 수백만 개의 홈 페이지 게시판에 등록된다. 또한, 하루에도 몇 번씩 올라오는 동일한 스팸 메시지는 사람이 직접 등록하는 것이 아니라 로봇 프로그램이 자동으로 등록하는 것이다. 인터넷 게시판에 등록되는 스팸 메시지는 사람이 직접 등록하거나 로봇을 이용하여 자동으로 등록한다[11].

(1) 사람이 직접 등록

사람이 인터넷 게시판에 직접 스팸 메시지를 등록하는 일반적인 과정은 (그림 1)과 같다. 인터넷 게시판에서 글쓰기 버튼을 클릭하여 글쓰기 화면의 양식에 따라 게시자의 이름과 이메일, 제목, 내용, 비밀번호, 자동글쓰기 방지기호 등을 입력하고, 등록하기 버튼을 클릭하면 등록된다.

(그림 1) 게시판의 메시지 등록 과정

(2) 로봇 프로그램을 이용한 등록

로봇 프로그램을 이용하여 인터넷 게시판에 스팸 메시지를 등록하는 방법은 사람이 직접 스팸 메시지를 등록하는 방법과 비교하여 다음과 같은 차이점을 가지고 있다.

첫째, 게시판 데이터베이스에 직접 데이터를 전송하여 스팸 메시지를 등록한다. 인터넷 게시판에 글을 작성하는 부분은 크게 2 단계로 나누어진다. 첫 번째 단계는 메시지를 입력하는 것이고, 두 번째 단계는 데이터를 데이터베이스에 등록하는 것이다. 로봇 프로그램은 데이터베이스에 직접 접속하여 데이터 정보를 바로 전송하여 메시지를 등록한다. 그렇기 때문에 단어 필터링, IP 차단 등의 방지 정책이 효과가 없다. 이렇게 글을 남기는 이유는 단어 필터링, IP 차단을 피해가려는 목적도 있지만 빠른 속도로 여러 개의 게시판에 메시지를 등록할 수 있기 때문이다.

둘째, 실제 사람이 접속하는 것과 같이 데이터를 생성하여 게시판 데이터베이스에 데이터를 전송하여 메시지를 등록한다. 실제 사람은 브라우저를 사용하여 인터넷 게시판에 접속한 후, 글쓰기를 눌러 메시지를 등록한다. 이 방식을 그대로 적용하여 사용하는 방식이다. 따라서 로봇 프로그램이지만 실제 사람이 입력하는 것과 거의 동일한 효과를 나타낼 수 있다. 하지만 단어 필터링, IP 차단 등으로 차단할 경우에는 메시지 등록을 막는 것이 가능하다.

셋째, 실제 사람이 메시지를 등록하는 것과 같이 데이터를 생성하여 게시판 데이터베이스에 데이터를 전송하여 메시지를 등록한다. 첫 번째와 두 번째의 장점을 결합하여 새롭게 등장한 방식이다. 실제 사람이 직접 입력하는 것과 같이 모든 정보(Cookie, Session, Referer, etc ...)를 가상으로 만들어 데이터베이스에 직접 전송한다. 따라서 이 프로그램으로 남겨지는 스팸 메시지는 단어 필터링, IP 차단 등이 통하지 않을 뿐만 아니라 빠른 속도로 여러 번 등록할 수 있는 기능이 있어 차단하기 어렵다.

2.2. 스팸 메시지 차단 방법

스팸 메시지 차단 방법에 관한 국내외의 연구가 미미한 실정이므로, 이 절에서는 유사한 스팸 메일 차단에 대한 대표적인 연구 결과인 Listing, 송신자 인증, 워드필터링(SpamAssassin, Bayesian Filtering) 등을 살펴본다.

(1) Listing에 의한 스팸 메시지 차단

게시자의 IP 주소, ID, 이메일 주소 등의 정보를 목록화하고 이에 부합되는 정보를 포함한 메시지를 스팸 메시지로 분류하는 방법이다. 이 방법은 높은 차단 효과를 위해 이미 알려진 스패머(Spamer)의 방대한 정보가 필수적이며, 이를 위해 스팸 메일 차단 정보는 전 세계 스패머의 정보를 공공 또는 사설 기관에서 수집하고, 이를 통해 스팸 메일의 차단을 시도하는 RBL(Real-time Black List)이 널리 이용되고 있다. 그러나 이 방법은 높은 차단 효과를 위해서 지속적으로 정보의 갱신과 관리를 필요로 하고, 메시지의 정보를 은닉하는 스패머의 간단한 트릭에도 유연히 대처하기 어렵다는 단점을 포함하고 있다[12, 13].

(2) 송신자 인증

스패머들이 메시지내의 정보를 은닉한다는 사실에 기반을 둔 것으로 메시지내의 송신자 정보가 실제 메시지의 송신자와 부합하는 가를 인증하는 방법이다. 이를 위해 스팸 메일 차단 분야는 이메일에 포함된 이메일 주소, 메시지를 송신한 IP 주소 등을 DNS(Domain Name Service)를 통해 인증하거나 이메일의 전자 서명을 첨부하여 파악하는 등의 기법들이 활용되고 있다. 이 방법은 최근 대형 정보기술업체와 ISP(Internet Service Provider)에 의해 발전되고 있으며, Pobox의 SPF[14], 마이크로소프트의 SenderID[15], 야후의 DomainKeys[16], 시스코의 IIM[17] 등의 방법들이 제시되고 있다. 그러나 게시판 스팸 차단 분야는 이메일 서비스와 달리 도메인 주소와 같은 공통된 기준이 없기 때문에 적용이 어렵다는 단점이 있다.

(3) 워드 필터링

Listing 방법과 같이 널리 이용되는 스팸 차단 방법으로 메시지에 포함된 단어 또는 문장을 바탕으로 스팸 메시지를 판단하는 방법이다. 이 방법은 규칙 기반 필터링 방법과 학습 기반 필터링 방법으로 구분된다.

SpamAssassin[18]으로 널리 알려진 규칙 기반 필터링 방법은 사전에 규칙(스팸 메시지에 포함된 단어나 문장)을 설정해두고, 이에 부합하는 단어나 문장을 포함한 메시지를 스팸 메시지로 간주하는 방법이다. 그러나 이 방법은 단어의 조합이나 표현을 변경하는 스패머의 간단한 트릭에도 유연하게 대처하기 어렵고, 높은 오탐지율을 나타낸다는 단점이 있다.

학습 기반 필터링 방법은 베이지안 필터링으로 널리 알려진 기법으로 과거의 메시지를 학습함으로써 새로 수신되거나 게재되는 메시지의 스팸 메시지 여부를 판단하는 방법이다. 이 방법은 개인의 학습에 따라 스팸 메일의 차단에는 효과가 좋으나 다수의 사용자에 대해 학습을 시키고 스팸 메시지를 차단하는 데에는 큰 효과를 발휘하지 못하는 단점이 있다.

3. 인터넷 게시판의 스팸 메시지 차단 방법

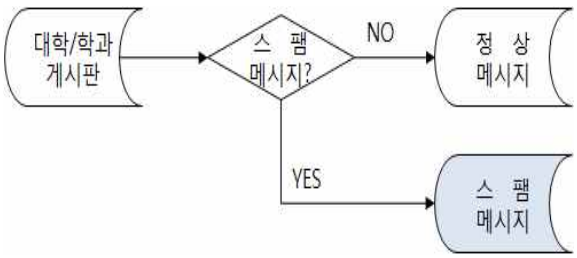
인터넷 게시판에 스팸 메시지를 남기는 기술이 갈수록 정교해지고 있기 때문에 스팸 메시지를 사전에 100% 식별하여 차단하는 것이 매우 어려운 현실이다. 따라서 이 장은 게시판 데이터베이스에 등록되는 스팸 메시지를 식별하여 차단하는 스팸 메시지 차단 방법을 설명한다. 제안한 방법은 일정 기간 동안 스팸 메시지를 수집하고, 수집한 스팸 메시지를 분석하여 정규화된 차단 규칙을 생성한 후, 이 규칙을 적용하여 게시판 데이터베이스에 등록되는 메시지에 대해 스팸 메시지의 여부를 식별하여 차단한다.

3.1 스팸 메시지 수집

인터넷 게시판의 성격에 따라 스팸 메시지를 서로 다르게 정의할 수 있기 때문에 홈 페이지 관리자의 정확한 스팸 메시지 식별에도 불구하고, 이를 수용하는 인터넷 이용자는

이를 잘못된 판단으로 인식하는 문제가 빈번하게 발생할 수 있다. 따라서 스팸 메시지를 특성에 따라 세분화시킴으로써 스팸 메시지의 정의에 대한 견해 차이를 해소하고, 인터넷 커뮤니티 상의 다양한 유형의 게시판이나 인터넷 이용자들의 요구사항을 충족시킬 수 있다.

본 논문에서 스팸 메시지(spam message)는 대학교 및 학과의 홈 페이지 게시판에 등록되는 메시지들 중 입시/입학, 수강, 등록 등 대학의 학사행정업무와 관련성이 없는 메시지들로 정의한다. 이러한 스팸 메시지 정의에 따라 대학교 및 학과의 홈 페이지 게시판에 등록되는 메시지들을 대상으로 (그림 2)와 같이 스팸 메시지를 식별하여 수집하였다. 그리고 메시지의 제목, 게시자, 이메일 등이 동일한 메시지가 2회 이상 반복적으로 등록될 경우에도 스팸 메시지로 분류하였다.



(그림 2) 스팸 메시지 수집 과정

<표 1>과 같이 수집된 모든 메시지는 897개이며, 정상 메시지 242개, 스팸 메시지 655개로 구성되어 있다. 수집된 메시지는 홈 페이지 게시판에 등록된 메시지이며, 등록된 메시지에 대한 등록일, 제목, 등록자, 등록자 이메일을 제외한 부가사항은 포함시키지 않았다.

〈표 1〉 스팸 메시지 수집 현황

수집기간	2009. 11. 1 ~ 2010. 2. 28 (4개월)
수 집 메 시 지	총 897개 메시지 (정상 메시지 242개, 스팸 메시지 655개)
수집 웹 사 이 트	대학교 및 학과의 홈 페이지 게시판

〈표 2〉 수집된 메시지의 구성과 스팸 메시지의 분류

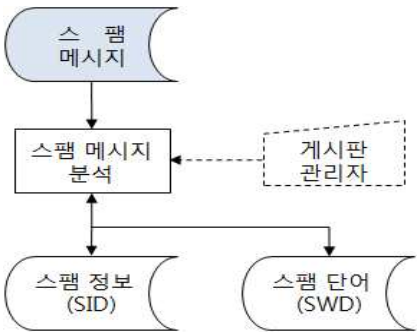
구분	정 상 메시지	스팸 메시지					계
		상업 광고	성인 광고	일반 광고	비방 욕설	기타	
수량	242	361 (55.11)	224 (34.20)	36 (5.50)	12 (1.83)	22 (3.36)	897
		655					
비율	26.97%	73.03%					100%

수집된 메시지는 <표 2>와 같이 먼저 정상 메시지와 스팸 메시지로 구분하며, 스팸 메시지는 상업(금융, 쇼핑, 판매 등) 광고, 성인(청소년유해, 음란, 성인용품, 게임, 경마 등) 광고, 건전성 일반 광고, 욕설 및 비방, 기타 메시지로 분류하여 스팸 메시지 차단의 정확성을 측정하는 과정에서 참조한다. 그리고 수집된 스팸 메시지의 소분류는 인터넷 게시판의 특성에 따라 변경될 수 있으며, 본 논문은 대학교 및 학과의 홈 페이지 게시판 특성에 한정하여 스팸 메시지를 분류하였다.

3.2 스팸 메시지 분석

스팸 메시지 분석은 대학교 및 학과의 홈 페이지 게시판에 등록되는 메시지로부터 수집한 스팸 메시지의 특성을 분석하여 스팸 메시지를 차단하기 위한 판단 자료로 이용할 수 있는 스팸 메시지 차단 규칙을 도출하기 위한 과정이다. 따라서 스팸 메시지를 분석하여 스팸 메시지 차단 규칙을 생성하는 과정에서 다음의 세 가지 목표를 달성할 수 있도록 고려하였다. 첫째, 스팸 메시지 차단에서 스팸 차단 비율을 높이기 위한 차단 규칙을 생성한다. 둘째, 새로운 유형의 스팸 메시지에 대해서도 판단이 가능한 차단 규칙을 생성한다. 셋째, 스팸 메시지 차단에서 판단 시간을 짧게 할 수 있는 판단 규칙을 생성한다.

위와 같이 스팸 메시지 차단 규칙의 생성에 반영되어야 하는 목표에 따라 스팸 메시지의 분석 및 차단 규칙 생성은 (그림 3)과 같이 진행하였다. 인터넷 게시판에 등록되는 메시지들 중에서 스팸 메시지의 정의에 따라 스팸 메시지를 분류하고, 스팸 메시지들을 대상으로 스팸 메시지 차단 과정에서 이용할 수 있는 스팸 정보 데이터베이스와 스팸 단어 데이터베이스를 생성한다.



(그림 3) 스팸 메시지 분석 및 차단 규칙 생성

스팸 정보 데이터베이스(SID : Spam Information Database)는 <표 3>과 같이 메시지 제목, 메시지 게시자 및 이메일, 메시지의 특성에 따른 분류, 그리고 동일 메시지의 등록 빈도수로 구성된다. 특히, 제목의 경우에는 단어 내 또는 단어 사이에 특수문자를 끼워 넣어서 스팸 메시지로 인식되지 않도록 변형된 단어를 이용하여 구성하는 경우가 많은데, 스팸 메시지 데이터베이스는 제목에서 이러한 유형의 특수문자를 제외시키고 구성하였다.

〈표 3〉 스팸 정보 데이터베이스(SID)

연번	제 목	게시자	이메일	분류	빈도
1	온라인경마 RACE.ES.PN 사설스크린경마,인터넷경마, ...	에스맨	nonoyse@yahoo.co.kr	상업	64
2	정품 비아그라 판매	필존	feelzoneus@yahoo.co.kr	성인	51
3	비아그라-필존!! 오픈2,000일 기념 이벤트, 할인 + 할인, ...	필존	feelzoneus@yahoo.co.kr	성인	34
4	엔진오일 최저가 공동구매 동호회 TNT클럽(순정부품, 튜닝용품, ...	TNT	sorry@sorry.com	상업	22
5	이미테이션 도매 홍콩 명품 도매 쇼핑몰 창업 상담	홍콩명품	twojobsmall@hotmail.com	상업	18
...					

스팸 단어 데이터베이스(SWD : Spam Word Database)는 <표 4>와 같이 스팸 메시지 제목에 포함된 스팸 단어와 빈도수로 구성되며, 단어사이에 특수문자를 끼워 넣어서 스팸 단어로 인식하지 못하도록 변형된 단어는 특수문자를 제외시키고 구성하였다. 스팸 단어는 스팸 메시지 가능성을 정확하게 판단하여 스팸 메시지 차단 비율을 높일 수 있는 규칙으로 사용된다.

〈표 4〉 스팸 단어 데이터베이스(SWD)

연번	스팸 단어	빈도
1	게임	98
2	라이브	96
3	비아그라	91
4	경마	83
5	다운	77
...	...	...

스팸 메시지의 수집과 분석 결과를 통해 “인터넷 게시판에 스팸 메시지를 올리는 광고주는 보통 하루에도 몇 번씩 스팸 메시지를 반복적으로 올리려고 시도하며, 필터링 시스템이 가동될 것을 예상하고 동일한 스팸 메시지를 반복적으로 올리거나 단어를 변형한 스팸 메시지를 지속적으로 올리려고 시도한다.”는 것이 확인되었다. 따라서 필터링 시스템을 한 번 적용한 것으로는 스팸 메시지 차단 효과가 없으며, 지속적으로 관리하며 갱신을 해주어야 비로소 효과를 볼 수 있다. 또한, 필터링 시스템의 스팸 메시지 차단 효과를 높이기 위해서는 스팸 메시지 차단 규칙을 정교하게 도출하여 적용해야 한다. 스팸 메시지 분석 결과로부터 생성된 스팸 정보와 스팸 단어를 이용하여 인터넷 게시판에 등록되는 스팸 메시지를 차단하기 위한 규칙을 <표 5>와 같이 도출하였다.

〈표 5〉 스팸 메시지 차단 규칙

순위	차단 규칙	참 조
1	게시되는 스팸 메시지 제목, 게시자, 이메일	SID
2	게시되는 스팸 메시지 제목 중 스팸 단어 및 빈도수	SWD
3	스팸 메시지의 게시자와 이메일	등록된 메시지
4	게시판 관리자	수동

첫 번째 규칙은 게시판에 등록되는 스팸 메시지의 제목, 게시자, 이메일을 이용하여 차단한다. 스팸 메시지의 90% 정도(표 2 참조)를 차지하는 상업 광고와 성인 광고 메시지의 대부분은 동일한 제목, 또는 특수문자를 제외시켰을 때 동일한 제목이 되는 메시지를 1회 이상 반복(표 3 참조)하여 등록하는 것으로 파악되었다. 또한, 반복하여 게시되는 스팸 메시지의 경우 동일한 게시자의 이름과 이메일을 이용하는 경우가 많은 것으로 파악되었다. 본 논문은 이러한 정보를 이용하여 스팸 메시지를 차단한다. 즉, 스팸 메시지로 분류되면 스팸 메시지의 제목, 게시자, 이메일, 소분류, 빈도수를 SID에 저장하고, 게시판 데이터베이스에 새로운 메시지가 등록되면 스팸 정보를 이용하여 다음과 같은 경우에 스팸 메시지로 식별하여 차단하게 된다.

- 게시판에 등록되는 메시지의 제목이 SID의 메시지 제목과 일치하면 차단
- 게시판에 등록되는 메시지의 제목이 SID의 제목과는 다르지만 게시자 이름 및 이메일이 SID의 게시자 이름 및 이메일과 일치하면 차단

두 번째 규칙은 게시판에 등록되는 스팸 메시지의 제목에서 추출한 스팸 단어를 이용하여 차단한다. 게시판에 스팸 메시지를 자동으로 등록시키는 프로그램을 이용하는 경우에는 제목, 게시자 이름 및 이메일 주소를 변형하여 스팸 메시지로 인식하지 못하도록 하는 경우가 있고, 이러한 경우는 첫 번째 규칙을 적용하여 차단할 수 없게 된다. 이러한 유형의 스팸 메시지를 차단하기 위한 방법으로 게시판에 등록된 메시지가 스팸 메시지로 분류되면 제목에서 단어를 추출하여 SWD를 구축하고, 이 정보를 이용하여 새롭게 등록되는 메시지에서 스팸 메시지를 식별하여 차단하게 된다.

- 새롭게 등록되는 메시지 제목에서 추출한 단어가 SWD에서 빈도수가 임계값(Θ)이상인 스팸 단어와 1개라도 일치하면 차단

한편, 임계값을 설정하는 방법은 여러 가지의 방법이 있을 수 있으나 본 논문에서는 스팸 단어들의 빈도에 대한 평균값을 구하는 방법을 사용하였다. SWD에 등록된 스팸 단어들의 빈도수 합을 구한 후 이를 스팸 단어의 개수로 나누

어서 평균 임계값을 구한다. SWD에 등록된 스팸 단어들의 빈도수를 $F_i(1 \le i \le n)$, 임계값을 Θ라고 하면 임계값은 다음 수식으로 구할 수 있다.

$$\Theta = \frac{\sum F_i}{n}, \ (1 \le i \le n)$$

세 번째 규칙은 게시판에 등록되는 스팸 메시지의 게시자 이름과 이메일 정보를 이용하여 차단한다. 첫 번째 규칙과 두 번째 규칙에 해당되지 않는 일부 스팸 메시지를 분석한 결과, 메시지를 게시하는 게시자 이름이나 이메일을 입력하지 않았으며, 게시자 이름이나 이메일을 입력하지 않는 메시지는 거의 대부분 스팸 메시지로 확인되었다. 따라서 게시판에 등록되는 메시지의 게시자 이름이나 이메일이 입력되지 않는 경우에는 스팸 메시지로 분류하여 차단하게 된다.

- 게시판에 등록되는 메시지의 게시자 이름 또는 이메일이 공백이면 차단

네 번째 규칙은 게시판 관리자가 지속적으로 관리하여 스팸 메시지를 식별하고 차단한다. 위의 세 가지 규칙에 해당되지 않은 메시지, 특히 <표 2>에서 스팸 메시지의 약 5.5%에 해당하는 일반 광고 메시지는 게시판 관리자가 판단하여 스팸 메시지 여부를 결정하여 차단해야 한다.

〈표 6〉 일반 광고 메시지의 예

제목	내용	게시자	이메일	분류	빈도
대한민국	오늘경기잘치르고내일좋은소식을들었으면좋겠습니다.	이다수	ado@daum.net	일반	1
1시간전	1시간전입니다.긴장되는군요,내가뛰는것도아닌데,꼭경기전의선수처럼,,,	대한	ioo@daum.net	일반	1

예를 들어, <표 6>과 같은 일반 광고 메시지는 도출한 스팸 메시지 차단 규칙으로 식별하는 것이 거의 불가능하다. 따라서 게시판 관리자가 메시지 내용을 확인하여 스팸 메시지 여부를 결정하고, 스팸 메시지로 분류되면 SID 및 SWD에 정보가 등록되어 동일하거나 유사한 메시지가 다시 등록될 경우에는 스팸 메시지 차단 규칙을 적용받게 된다.

3.3 스팸 메시지 차단

스팸 메시지 분석에서 만들어진 스팸 메시지 차단 규칙을 이용하여 인터넷 게시판에 등록되는 스팸 메시지를 차단한다. 스팸 메시지 차단 과정에서 스팸 메시지 여부를 검사하는 시간을 단축하기 위해 수행시간이 빠른 검사 방법을 먼저 수행하며, 스팸 메시지를 식별하기 위한 검사 순서 및 방법은 다음과 같다.

(1) 스팸 메시지의 정보 검사

비교 정보 판별규칙을 이용하여 스팸 메시지를 검사하는 것으로 판별 규칙에서 스팸 메시지의 제목이 같은 것이 있는지 검사하며, 스팸 메시지를 게시하는 게시자 이름과 이메일이 같은 판별규칙이 존재하는지 검사한다.

(2) 스팸 단어의 빈도 검사

스팸 메시지의 제목에 포함된 스팸 단어의 빈도 판별 규칙은 제목에 포함된 각 스팸 단어에 대해 출현 빈도가 임계값(Θ)이상인 단어를 1개라도 포함하고 있는 판별 규칙이 존재하는지 검사한다.

(3) 스팸 메시지의 게시자 정보 검사

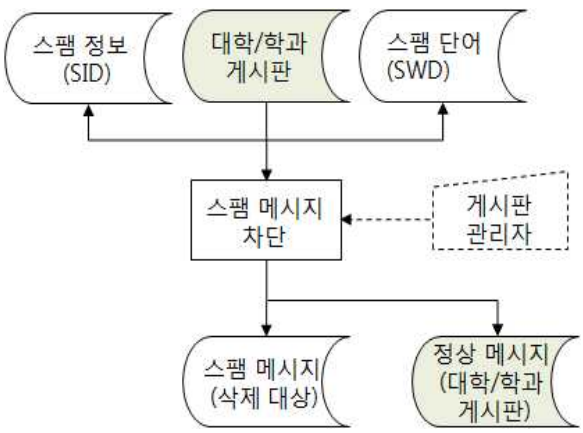
대부분의 인터넷 게시판에서 메시지를 게시할 때, 게시자의 이름과 이메일 주소를 입력하도록 되어 있다. 그러나 스팸 메시지를 게시판 데이터베이스에 자동으로 등록시키는 경우에는 게시자의 이름이나 이메일 주소를 입력하지 않은 경우가 종종 있다. 이러한 경우에는 스팸 메시지로 처리하기 위해서 메시지 게시자의 이름과 이메일 주소를 검사한다.

4. 구현 및 스팸 메시지 차단의 정확도 측정

이 장은 스팸 메시지의 수집과 빈도 분석을 이용한 스팸 메시지 차단 시스템을 구현하고, 구현된 시스템을 실제 인터넷 게시판 환경에서 운영한 결과를 바탕으로 스팸 메시지 차단의 정확도를 측정한 실험 결과를 기술한다.

4.1 스팸 메시지 차단 시스템의 구현

인터넷 게시판에 등록되는 스팸 메시지를 차단하는 시스템은 (그림 4)와 같이 구성되었다. 게시판에 메시지가 등록되면 SID 및 SWD를 바탕으로 스팸 메시지 차단 규칙에 따라 스팸 메시지 여부를 식별하여, 스팸 메시지로 식별되면 SID 및 SWD를 갱신하고 해당 스팸 메시지를 삭제한다. 스팸 메시지 차단 시스템은 인터넷 게시판에 등록되는 메시지를 대상으로 정상 메시지와 스팸 메시지를 식별하는 과정에서 이미 차단된 스팸 메시지의 정보를 활용한다.



(그림 4) 스팸 메시지 차단 시스템의 구성도

스팸 메시지의 판단 과정에서 스팸 메시지 판단 규칙을 반복 적용함으로써 스팸 메시지 여부를 판단하게 된다. 예를 들어, 인터넷 게시판에 임의의 메시지가 등록되면, ① 해당 메시지에 대해 제목이 SID의 제목과 일치하는지 여부, ② 해당 메시지의 게시자 및 이메일이 SID의 게시자 및 이메일과 일치하는지 여부, ③ 해당 메시지의 제목에 포함된 단어가 SWD에 포함된 스팸 단어(빈도수>θ)와 일치하는지 여부, ④ 해당 메시지의 게시자 또는 이메일이 입력되지 않았는지 여부를 차례로 판단하게 된다. 새로 등록된 메시지가 이 가운데 하나의 차단 규칙에라도 해당하면, 이 메시지는 스팸 메시지로 식별하여 삭제하며, 삭제되는 스팸 메시지에 대한 정보는 SID 및 SWD에 추가하거나 갱신한다.

4.2 정확도 측정 환경

제안한 스팸 메시지 차단 시스템의 정확성을 검증하기 위하여 (그림 4)와 같이 스팸 메시지 차단 시스템을 구현하였으며, 대학교 및 학과의 홈 페이지 게시판에 적용하여 실험하였다.

스팸 메시지 차단 시스템의 구현과 실험에 사용된 시스템 환경 및 구현 도구 등의 실험 환경은 <표 7>과 같다. 스팸 메시지 차단 시스템은 Windows XP/Vista 환경에서 Visual C++을 이용하였고, 스팸 정보와 스팸 단어의 데이터베이스는 MS-SQL을 이용하여 구현하였다.

〈표 7〉 구현 및 실험 환경

구 분		명 칭
게시판 DB	OS	UNIX(Linux), Windows 2003 server
	DB	Oracle, MS-SQL 2003
	웹 서버	Apache, IIS
스팸 차단 시스템	OS	Windows XP or Vista
	DB	MS-SQL 2003
	구현 언어	Visual C++

한편, 게시판은 누구나 접속하여 제목, 내용, 게시자, 이메일 정보만 입력하면 메시지를 등록할 수 있도록 구성하였고, 로봇 프로그램을 이용하여 게시판 데이터베이스에 메시지를 자동으로 등록하는 것도 허용하였다. 인터넷 게시판에 등록되는 메시지들을 대상으로 스팸 메시지 차단의 정확성을 측정하기 위해 실험에 사용한 메시지의 구성은 <표 8>과 같다.

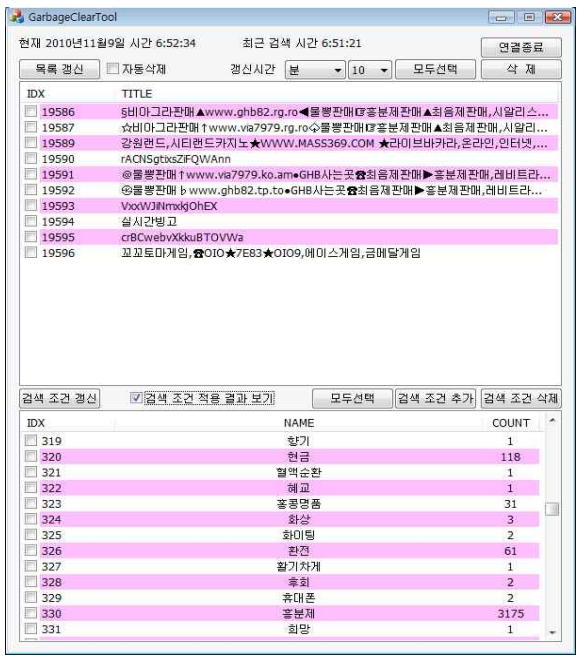
〈표 8〉 실험에 사용된 메시지의 구성

실험기간	2010. 3. 1 ~ 2010. 8. 31 (6개월)
실험대상 메 시 지	총 1,658개 메시지 (정상 메시지 385개, 스팸 메시지 1,273개)
실험대상 사 이 트	대학교 및 학과의 홈 페이지 게시판

4.3 정확도 측정 결과

게시판 데이터베이스에 새롭게 등록되는 메시지들을 대상(그림 5)으로 스팸 메시지 여부를 식별하여 차단하는 시스템의 스팸 메시지 식별 및 차단 모니터링 화면은 (그림 6)과 같다. 게시판 데이터베이스에 메시지가 등록되면 실시간 또는 주기적으로 게시판 데이터베이스를 확인하여 스팸 메시지 판별 규칙을 적용하여 스팸 메시지로 식별되면 스팸 메시지의 제목, 게시자의 이름과 이메일 정보는 SID, 스팸 단어는 SWD에 등록 또는 갱신한 후 해당 메시지를 삭제한다.

(그림 5) 스팸 메시지가 등록된 게시판 화면



(그림 6) 스팸 메시지 식별/차단 모니터링 화면

(그림 6)과 같은 스팸 메시지 차단 시스템을 이용하여 인터넷 게시판에 등록되는 메시지를 대상으로 제안한 스팸 메시지 차단 방법을 적용하여 스팸 메시지 차단의 정확도를 측정하였다. 정확도는 정상 메시지와 스팸 메시지의 단순 분류와 스팸 메시지를 5개 영역으로 세분화하여 측정하였다. <표 9>은 스팸 메시지의 정의에 따라 정상 메시지와 스팸 메시지로 분류했을 때의 스팸 메시지 식별의 정확도를 나타내며, <표 10>는 스팸 메시지를 5개의 영역으로 세분화했을 때의 스팸 메시지 차단의 정확도에 실험 결과를 나타낸다.

〈표 9〉 스팸 메시지 식별 실험 결과

전체 메시지	정상 메시지			스팸 메시지		
	실제 메시지	판단 메시지	오탐지	실제 메시지	판단 메시지	오탐지
1,658	385 *(23.2)*	374 (97.1)	11 (2.9)	1,273 *(76.8)*	1,222 (96.0)	51 (4.0)

〈표 10〉 스팸 메시지 차단의 정확도

분 야	실제 스팸 메시지	차단 스팸 메시지	차단율	오탐지율
상업 광고	534(41.9%)	523	97.9	2.1
성인 광고	610(47.9%)	601	98.5	1.5
일반 광고	71(5.6%)	48	67.6	32.4
비방/욕설	27(2.1%)	23	85.2	14.8
기 타	31(2.4%)	27	87.1	12.9
합 계	1,273	1,222	96.0	4.0

<표 9>의 실험 결과를 살펴보면 홈 페이지 게시판에 등록되는 전체 메시지의 약 76% 이상이 스팸 메시지로 식별되었으며, 정상 메시지는 385개 중 374개를 식별하여 97.1%의 정확도, 스팸 메시지는 1,273개 중 1,222개를 식별하여 96.0%의 정확도를 나타냈다. 즉, 정상 메시지를 스팸 메시지로 식별하는 오탐지율은 2.9%, 스팸 메시지를 정상 메시지로 식별하는 오탐지율은 4.0%로 비교적 낮게 나타났다.

<표 10>은 스팸 메시지를 메시지의 특성에 따라 5개 영역으로 세분화하여 정확도를 측정하였다. 스팸 메시지의 약 90% 이상을 차지하는 상업 광고와 성인 광고 메시지의 차단율이 각각 97.9%, 98.5%로 매우 높게 나타났으나 일반적인 광고 메시지의 차단율은 67.6%로 상대적으로 낮게 나타나고 있다. 상업 광고나 성인 광고 메시지는 동일한 제목 또는 동일한 게시자의 이름 및 이메일로 반복해서 등록하거나 메시지 제목에 스팸 메시지에 자주 등장하는 스팸 단어가 포함되기 때문에 본 논문에서 제안한 스팸 메시지 차단 규칙에 의해 대부분 차단되는 것으로 판단된다. 그러나 일반적인 광고 메시지는 한 번만 등록하거나 게시판의 성격에 따라 스팸 메시지로 판단하기 어려운 경우가 많기 때문에 본 논문에서 제안한 스팸 메시지 차단 규칙으로 차단이 어려운 것으로 판단되며, 이에 대한 원인 분석 및 해결 방안이 도출되어야 할 것으로 생각한다.

4.4 기존 방법과의 비교

인터넷 게시판에 등록되는 스팸 메시지를 차단하기 위한 기존의 연구 결과가 거의 없기 때문에 스팸 메일 차단 방법과 스팸 메시지 차단 방법에 대한 기존 연구 결과를 <표 11>과 같이 제안한 방법과 비교하였다. 제안한 방법은 인터넷 게시판에 등록되는 메시지들을 대상으로 스팸 메시지를 식별하여 차단하는데 필요한 정보를 획득하기 위해 제안한

〈표 11〉 스팸 메일 및 스팸 메시지 차단 방법의 비교

항목 \ 방법	스팸 메일 차단 방법			스팸 메시지 차단 방법	
	백기영 외[2]	RBL	SpamAssassin	김범배 외[1]	제안한 방법
스팸 정보 수집	자동 수집	사용자 등록	사용자 판별에 의한 Learning	사용자 판별에 의한 Learning	자동 수집 + 사용자 등록
가중치 적용	빈도 + 시간	X	빈도(단어)	빈도(단어)	빈도(제목,단어)
스팸 정보 공유 가능	가능	가능	서버별	서버별	가능
스팸 정보 분석 시간	중간	빠름	느림(본문분석)	느림(본문분석)	빠름
서버 부하	중간	낮음	높음(본문분석)	높음(본문분석)	중간
차단 방법의 다양성	다양한 방법	보낸 메일서버 IP	빈도 조사	빈도 조사	다양한 방법
스팸 탐지/ 차단의 정확성	높음	보통	보통	보통	높음
오탐지율 (일반, 스팸)	낮음	보통	높음	높음	낮음

방법으로 식별 및 차단한 스팸 메시지를 이용한다. 즉, 스팸 메시지의 식별 및 차단을 위해 스팸 메시지를 이용하기 때문에 기존의 방법들에 비해 스팸 메시지 탐지 및 차단의 정확성이 높고, 다양한 유형의 스팸 메시지나 진화하는 스팸 메시지에도 능동적으로 대응할 수 있다. 또한 정상 메시지에 대한 오탐지율이 2.9%, 스팸 메시지에 대한 오탐지율이 4.0%로 기존의 방법들에 비해 상대적으로 낮은 것으로 확인되었다.

5. 결 론

인터넷 게시판에 등록되는 모든 스팸 메시지를 100% 차단하는 것은 사실상 불가능하다. 스팸 메시지를 남기려는 사람과 차단하려는 사람의 치열한 공방전이 일어나고 있고, 프로그램의 지능과 패턴이 갈수록 정교해지고 있다. 스팸 메시지 차단 정책을 최대한 적용해도 어느 정도의 스팸 메시지는 올라올 수밖에 없는데, 이를 전략적으로 관리하지 않고, 보이는 대로 삭제하면 인력과 시간의 낭비가 클 뿐만 아니라 홈 페이지에 방문하는 고객, 이용자들에게도 나쁜 영향을 끼쳐 결국 신뢰도 없는 죽은 홈 페이지가 될 수 있다.

본 논문은 인터넷 게시판에 등록되는 스팸 메시지를 수집하여 메시지의 특성 및 빈도를 분석하고 차단하는 단계로 구성된 게시판 스팸 메시지 차단 방법을 제안하였다. 인터넷 게시판의 데이터베이스에 저장되는 모든 메시지를 대상으로 스팸 메시지를 수집하고, 수집한 스팸 메시지를 분석하여 스팸 메시지를 식별할 수 있는 정규화된 차단 규칙을 생성한 후, 이 규칙을 이용하여 등록된 메시지에 대해 스팸 메시지를 검사하여 차단한다.

제안한 스팸 메시지 차단 방법은 스팸 메시지의 대부분을 차지하고 있는 상업 광고와 성인 광고 관련 스팸 메시지의 경우는 98% 이상 차단이 가능한 것으로 확인되었다. 다만, 게시판의 특성에 따라 스팸 메시지로 분류가 애매모호한 일반적인 광고 메시지의 차단율이 상대적으로 낮게 나타나고 있는 부분은 해결되어야 할 과제로 생각한다.

본 논문에서 제안한 스팸 메시지 차단 방법은 실험 결과와 같이 다양한 유형의 스팸 메시지를 식별하여 차단할 수 있으며, 날로 발전하고 있는 스팸 메시지의 형태에 대해서도 유연하게 대처할 수 있는 구조를 가지고 있다.

그러나 인터넷 게시판에 적합한 메시지로 위장하여 일반적인 광고 메시지를 남기거나 로봇 프로그램을 이용하는 등 지능화된 기법을 사용하여 스팸 메시지 차단을 점점 더 어렵게 만들고 있다. 따라서 일반적인 광고 메시지를 식별할 수 있는 차단 규칙 및 모든 유형의 인터넷 게시판에 적용할 수 있도록 일반화된 차단 규칙에 대한 연구를 수행할 예정이다.

참 고 문 헌

[1] 김범배・최형기, 베이지안을 이용한 인터넷 커뮤니티 상의 유해 메시지 차단 기법, 정보처리학회논문지C, Vol.13-C, No.6, pp.733~740, 2006.

[2] 백기영・김승해・최장원・류재철, 수집과 빈도분석을 통한 스팸 메일 차단 방법, 정보처리학회논문지C, Vol.12-C, No.1, pp.137~146, 2006.

[3] Paulson, L.D., Spam Hits Instant Messaging, IEEE Computer, Vol.37, No.4, pp.18~18, 2004.

[4] 공미경・이경순, 스팸성 자질과 URL 자질의 공동 학습을 이용한 최대 엔트로피 기반 스팸 메일 필터 시스템, 정보처리학회논문지B, Vol.15-B, No.1, pp.61~68, 2008.

[5] 김명원・곽후근, 튜링 테스트 기반으로 한 VoIP 스팸 방지, 정보과학회논문지 : 컴퓨팅의 실제 및 레터, Vol.14, No.3, pp.261~265, 2008.

[6] 김범배・최형기, 신경망과 유전자 알고리즘을 이용한 스팸 메일 필터링 기법의 구현과 성능평가, 정보처리학회논문지C, Vol.13-C, No.2, pp.259~266, 2006.

[7] 김종민・김형근・김봉기, 스팸 메일 차단을 위한 RBL 개념의 확장에 관한 연구, 한국해양정보통신학회논문지, Vol.12, No.10, pp.1808~1814, 2008.

[8] 백기영・이철수・류재철, URL 빈도 분석을 이용한 스팸 메일 차단 방법, 정보보호학회논문지, Vol.14, No.6, pp.135~148, 2004.

[9] 이신영・길아라・김명원, 링크 구조 분석을 이용한 스팸 메일 분류, 정보과학회논문지 : 소프트웨어 및 응용, Vol.34, No.1, pp.30~39, 2007.

[10] 조인휘・심혜택, 휴대폰 SMS를 위한 SVM 기반의 스팸 필터링 시스템, 한국통신학회논문지, Vo.34, No.9, pp.908~913.

[11] 게시판 스팸차단 연구소, 게시판 스팸 차단 노하우 : 게시판 스팸을 차단하는 11가지 비법, http://cleanboard.net.

[12] SpamCop, SpamCop Blocking List, http://www.spamcop.net/bl.html.

[13] Spamhaus, The Spamhaus Block List, http://www.spamhaus.org/sbl/index.lasso.

[14] Pobox, SPF, How it works, http://넬.pobox.xom/howworks.html.

[15] Microsoft SenderID, Sender ID Framework Overview, http://www.microsoft.com/mscorp/safety/technologies/senderid/overview.mspx.

[16] Yahoo! DomainKeys, DomainKeys : Proving and Protecting Email Sender Identity, http://antispam.yahoo.com/domainkey.

[17] Jim Fenton, Identified Internet Mail, Cisco System, https://antiphishing.kavi.com/events/ Conference_Notes/ Jim_Fenton_on_Cisco_Internet_Identified_Mail.pdf.

[18] Graham Paul, A Plan for Spam, http://www.paulgraham.com/spam.html, 2002.

[19] SpamAssassin, The Apache SpamAssassin Project, http://spamassassin.apache.org.

김 태 희

e-mail : thkim@dsu.ac.kr

1991년 동신대학교 전자계산학과(공학사)

1993년 전남대학교 전산통계학과 (이학석사)

1999년 전남대학교 전산통계학과 (이학박사)

1997년~현 재 동신대학교 디지털콘텐츠학과 부교수

관심분야 : 소프트웨어공학, 객체지향 모델링, 컴퓨터 교육

강 문 설

e-mail : mskang@gwangju.ac.kr

1986년 전남대학교 전산통계학과(이학사)

1989년 전남대학교 전산통계학과 (이학석사)

1994년 전남대학교 전산통계학과 (이학박사)

1994년~현 재 광주대학교 컴퓨터공학과 교수

관심분야 : 소프트웨어공학, 정보보호관리, 인터넷 윤리, 컴퓨터 교육