http://dx.doi.org/10.5392/JKCA.2011.11.12.122

사용자 맞춤형 스팸 문자 필터링 시스템

Personalized Mobile Junk Message Filtering System

이승재, 최덕재
전남대학교 전자컴퓨터공학과

Seung-Jae Lee(lsj8902@paran.com), Deok-Jai Choi(dchoi@jnu.ac.kr)

요약

스팸 문자 메시지는 모바일 이용자에게 불쾌감을 줄 뿐만 아니라 불필요한 사회비용을 유발하는 유해 요소이다. 특히 스마트워크 시스템에서 핵심 단말인 스마트폰으로 유입되는 스팸 문자는 업무능률 향상이라는 스마트워크의 취지를 무색하게 만들 수 있어 이에 대한 연구가 필요하다. 본 논문에서는 스팸 자동분류기로 스팸 메시지를 차단함에 있어서, 오분류 결과를 학습군에 재반영하여 연산량을 줄이고 인식 성능을 개선할 수 있는 방법을 제안하였다. 스팸 분류기는 스마트폰에서 독립적으로 동작하고, 사용자의 수신 메시지만으로 학습하므로 사용자의 분류 판단 성향을 반영할 수 있다. 많은 컴퓨팅 자원을 소비해야 하는 전처리, 특징 선정, 훈련 과정은 사용자의 인증 컴퓨터가 담당하고 필터링 과정만을 스마트폰에서 처리한다. 실험 결과 95%이상의 양호한 결과를 보였고 스팸 분류기는 스마트폰의 일정 자원만을 점유하면서 동작하였다.

■ 중심어 : | 스팸 문자 메시지 | 스마트 폰 | 스팸 필터 |

Abstract

Mobile spam message is a harmful factor which makes receivers to be annoyed and leads to unnecessary social cost. Unwanted junk messages flowing to a smart phone ruin main purpose of the smart work system to enhance the productivity, so we need to study on this area. In this paper, we proposed a novel spam filter on the smartphone in order to reduce computing process and improve the accuracy rate by feedback of error results to a training sample set. As the spam classifier operates on the smartphone independently with training on only user's received data, it could reflect user preference. The authorized personal computer takes on heavy works, such as preprocessing, feature selecting and training process, and the smartphone takes on light works to block junk messages. Experimental results showed reasonable accuracy rate of over 95%, and we found that the application occupied constant computing resources while running on the phone.

■ keyword : | Mobile Spam Message | Smartphone | Spam Filter |

* 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2011-C1090-1111-0008)

접수번호 : #111017-004
접수일자 : 2011년 10월 17일
심사완료일 : 2011년 11월 22일
교신저자 : 이승재, e-mail : lsj8902@paran.com

Ⅰ. 서 론

이동통신 기술의 발전으로 모바일 기기는 우리 생활에 없어서는 안 될 필수품이 되었다. 일반 사용자들은 모바일 SMS(Short Message Service)를 개인 통신 수단으로 이용하고, 각종 기관, 단체, 모임 주체들은 SMS를 공지, 긴급 통신 수단 등으로 이용하고 있다. 하지만 SMS는 주요한 정보 전달 수단인 동시에 원치 않는 스팸 정보의 전달 매체로 악용되기도 한다. KISA(Korea Internet Security Agency)에 따르면, '불법 스팸대응센터'에 접수된 SMS 스팸 신고 접수 건수가 해마다 큰 폭으로 증가하고 있다. 스팸 메시지는 수신자에게 불편과 짜증 등 정신적인 피해를 야기하며 시간낭비와 생산성 감소 등 불필요한 사회비용을 유발시킨다. 나아가 이는 정부가 실시하고 있는 스마트워크 프로젝트[1]의 '생산성 향상'이라는 취지를 퇴색시킬 수 있다. 스마트워크 시스템의 핵심인 모바일기기로 불필요한 SMS 스팸 데이터가 자주 인입 되면 근무자의 작업 집중력을 저하시키고 SMS 정보의 신뢰성을 떨어뜨린다. 따라서 스마트기기 유저에게 모바일 스팸 메시지 차단은 반드시 필요하다.

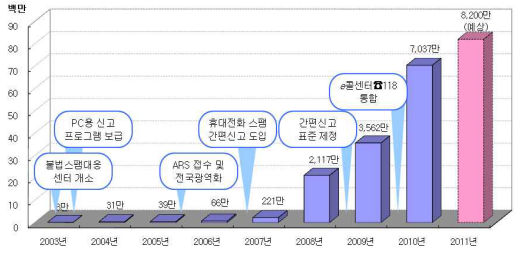


그림 1. 연도별 스팸신고 접수 증가 추이[출처:KISA]

스팸 문자 메시지를 자동 차단하는 기법은 크게 '단순 키워드 매칭' 방식과 '문서 자동 분류' 방식으로 구분할 수 있다. 단순 키워드 매칭 방식은 구현하기가 매우 쉽고 컴퓨팅 자원 소모가 매우 적지만, 사용자가 특정 번호, 문구 등을 직접 입력해야 하므로 사용자 편의성이 매우 낮다. 게다가 키워드 매칭은 'ㅂrㄷr', 'ㅇI야ㄱI', 'vi@gra', 'b00k', 'B11L'와 같은 스팸 키워드에 대한 의도적 변형에 매우 취약하여 인식율이 매우 낮다. 이에 반해 확률 모델이나, 인공지능 알고리즘을 사용하는 '문서 자동 분류' 방식은 키워드 입력에 대한 불편함이 없고 인식율도 양호하다. 하지만 처리 과정이 복잡하고 많은 컴퓨팅 자원을 소모하므로 모바일 기기가 모든 과정을 처리하는 것은 어렵다.

본 연구에서는 사용자의 개인 성향을 반영하기 위하여 사용자의 휴대단말기에서 수신한 메시지를 기반으로 학습 및 업데이트가 가능한 필터 모델을 개발하고자 한다.

본 연구에서 개발한 기술을 이용할 경우 사용자가 손쉽게 스팸차단을 할 수 있을 뿐만 아니라, 필터링으로 인한 부가적으로 발생되는 트래픽이 없어서 유무선 네트워크에도 부하를 주지 않을 것으로 기대된다.

본 논문의 I장 서론은 모바일 스팸 메시지의 실태에 대해 설명하고 분문 II장 관련연구에서는 기존의 SMS 스팸차단 기술과 그 기술들의 약점을 지적한다. III장은 자동 문서분류에 관련된 이론과 제안하는 스팸 필터링 시스템에 대한 구성 및 특징에 대해서 설명한다. IV장은 시스템의 설계 및 구현에 관하여 기술하였고 V장은 샘플 데이터에 대한 설명과 실험 결과 및 분석을 정리하였다. 마지막 VI장에서는 본 연구 결과를 정리하고 향후 연구 과제를 제시하였다.

Ⅱ. 관련연구

1. SMS 스팸

스팸이란 수신자의 의사에 반하여 정보통신망을 통해 일방적으로 전송되는 영리목적의 광고성 정보를 말한다. 스팸은 수신자에게 불편과 짜증 등 정신적인 피해를 야기하며 시간낭비와 생산성 감소 등의 불필요한 사회비용을 유발할 뿐만 아니라 메시지 전송·저장에 따른 네트워크 자원 소모를 가중시켜 정보통신서비스 제공자에게 비용을 전가하고 ICT(Information and Communication Technology) 녹색성장을 저해한다.

이에 정부 주도하에 스팸 근절을 위한 여러 가지 시도가 지속되고 있으며, 주요한 스팸 차단 방법은 아래와 같이 요약할 수 있다[2].

- 이동통신사의 스팸차단 서비스 개발 유도
- 휴대 단말기에 스팸관리기능 탑재 유도
- 스팸 신고에 의한 법적 규제

하지만, 위의 대응 방법에도 불구하고 사용자 설정에 대한 번거로움과 지능화된 스팸 전송수법 등에 의하여 스팸은 통제되지 않고 있다.

SMS 전달 경로에서 스팸 메시지를 차단할 수 있는 지점은 아래 [그림 2]와 같이 송신측 SMS 센터, 수신측 SMS 센터, 수신자 단말기이다. 그리고 각 지점에서 스팸문자를 차단하는 기법은 아래와 같이 정리할 수 있다.

- 요청 응답에 의한 차단
- 문자 송신 행태로 구분 차단
- 가명 사용에 의한 차단
- 블랙리스트에 의한 차단
- 문서 분류 기법에 의한 차단

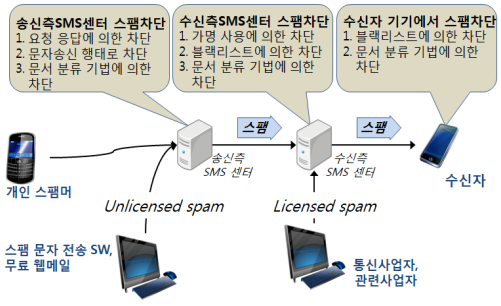


그림 2. 스팸 문자 메시지 차단 방법

2. SMS 스팸 차단 기법

2.1 요청 응답에 의한 차단 (Call and Response)

이 방법은 송신자 측의 셀룰러 네트워크에 적용 가능한 방법으로서, 특정 소프트웨어 프로그램을 이용하여 무선 네트워크에 스팸 메시지를 자동으로 대량 유입하는 경우를 예방할 수 있는 방식이다. 송신자가 메시지를 전송하면 SMS 센터는 다시 송신자에게 *확인 요청* 메시지를 보내고 송신자는 이에 대한 적절한 *확인 응답*을 SMS 센터에 되돌려 준다. 이 방법은 지능을 가진 사람은 쉽게 인식하여 응답할 수 있지만, 지능이 없는 컴퓨터는 응답할 수 없는 일종의 튜링 테스트이다[3]. 송신자가 SMS 센터에 보낸 *확인 응답*이 적절하면 목적지 네트워크로 포워딩 되지만 적절하지 못한 메시지는 스팸으로 간주되어 차단된다.

2.2 송신 행태로 구분 차단 (Sending-Behavior-Based)

스팸 메시지를 보내는 사람들은 가능한 여러 사람들에게 무작위로 대량의 메시지를 동시에 전송하고자 한다. Hu와 Yan[4]은 시간 영역에서 빈도를 사용하여 스팸 메시지를 구분하였다. 이들은 스팸 메시지 전송 행태를 파악하기 위하여 SMS 전송 시스템의 시스템 로그를 수집하고 시간-빈도 영역에서 이들을 분석하였다. 사전 데이터를 바탕으로 스팸 전송의 패턴을 유추하여 대량 전송되는 SMS의 스팸 여부를 판단할 수 있었다.

반면에 Qian[5]은 ID3(Interactive Dichotomizer 3)를 이용하여 스팸 또는 햄 메시지(정상 메시지)의 전송 행태를 구분하였다. 전송 행태로서 같은 메시지를 지속적으로 얼마나 보내는지 그리고 수신자까지 전송 성공률은 얼마나 되는지를 측정하였다. 전송 행태에 의한 차단 방식은 SMS 센터에 적용되는데 메시지가 도착하면 기존 학습 데이터를 기반으로 스팸인지 아닌지의 가능성을 탐색한다. 이 방법은 센터에서 스팸을 미리 차단함으로써 불필요한 무선 접속망의 대역폭 낭비를 줄일 수 있다. 하지만 이 방법을 적용하기 위해서는 여러 시스템으로부터 가능한 많은 양의 데이터 수집이 필요하다. 설사 충분한 데이터를 수집한다 하더라도 스팸을 전송하는 행태는 매번 또는 의도적으로 변화 될 수 있으므로 학습 데이터 수집은 계속 유지되어야 한다.

2.3 앨리어싱에 의한 차단 (Aliasing)

이 방법은 AT&T(American Telephone and Telegraph) 통신사처럼 이메일을 통한 SMS 전송 서비스를 제공하는 경우에 해당되는 스팸 메시지 차단 방식이다. AT&T 가입자에게 [phone number]@txt.att.net와 같이 이메일을 전송하면 가입자에게 SMS로 전송된다. 이 서비스를 이용하여 메시지를 전송하는 송신자는 요금을 지불하지 않아도 되므로, SMS 이메일 서비스는 스팸 전송에 악의적으로 이용되었다.

이러한 문제점을 해결하기 위해 전화번호 대신에 가명(alias)을 사용하여 이 서비스를 제공하였다. 즉 [phone number]@txt.att.net으로 전송하면 차단되고 [alias name]@txt.att.net으로 전송해야 가입자에게 SMS로 전달된다. 이 방식은 무료 이메일을 이용하여 대량 스팸 메시지를 악의적으로 전송하는 스팸머들을 방지 할 수 있지만, 전화번호 이외의 가명(alias name)을 별도로 알려야 하는 번거로움이 있다.

2.4 블랙리스트에 의한 차단(Black and White List)

블랙리스트에 의한 차단 방식은 현재 여러 모바일 기기에 탑재되어 보편화 되어 있는 방식이다. 사용자는 송신자의 전화번호나 스팸 키워드로 직접 블랙리스트를 만들고 이를 필터링 시스템에 적용한다. 새로 인입되는 메시지마다 전화번호, 키워드를 검사하여 일치하는 항목이 있으면 사용자에게 문자 도착 사실을 알리지 않고 자동으로 내부의 스팸 박스에 저장한다. 반면 사용자의 휴대폰에 저장되어 있는 전화번호는 화이트리스트로서, 화이트리스트 SMS가 도착하면 햄으로 간주되어 분류된다. 이 방식의 장점은 모든 문자 메시지를 수신 저장하므로 오분류된 결과가 있는지 사용자가 확인할 수 있다는 점이다. 이 방식은 얼마나 양호한 블랙리스트를 작성하느냐에 따라 필터링 결과에 큰 차이가 나타날 수 있다. 그리고 블랙리스트를 사용자가 직접 작성해야 하는 불편한 점도 있다.

2.5 자동문서 분류에 의한 차단(Text Classification)

문서 분류 기법에 의한 차단 방식은 다른 방식에 비해 적은 비용으로 사용자의 편의성을 높일 수 있는 방법이다. 이 방식은 SMS의 내용을 기반으로 스팸과 햄을 구분하는데, 샘플 데이터를 이용하여 필터링 시스템을 학습시키는 과정이 필요하다는 것이 다른 방식과의 큰 차이점이다. 문서 분류 기법을 이용하는 대부분의 필터링 시스템은 학습데이터 수집, 특징 단어 추출, 특징 단어 선정, 벡터 생성, 학습 및 인식 과정을 거친다. 문서 분류 기법은 사용 언어, 문자에 따라 불용어 제거, 특징 단어 추출 등의 전처리 기법이 달라지고, 학습 과정에 사용하는 데이터 샘플이 얼마나 양호하느냐에 따라 필터링 성능이 좌우되기도 한다. 따라서 학습용 샘플 데이터는 스팸, 햄 집합을 확연하게 구분할 수 있는 양질의 데이터를 가능한 많이 확보해야 한다.

3. 기존 SMS 스팸방지 시스템의 약점

기존 스팸 차단 기법들은 [표 1]과 같이 요약 비교할 수 있고 주요한 약점들은 다음과 같이 정리할 수 있다.

- 사용자 개인 성향 반영의 어려움
- 오분류 결과에 대한 업데이트 어려움
- 오분류 결과에 대한 책임소재
- 스팸필터 동작에 따른 부가 트래픽 유발
- 문서분류의 경우 많은 연산량에 대한 부담

표 1. 스팸차단 기법 비교

기준 \ 기법	요청응답 2.1	송신행태 2.2	가명사용 2.3	블랙리스트 2.4	문서분류 2.5
필터링 위치	송신 SMS센터	송신 SMS센터	수신 SMS센터	수신 SMS센터, 단말	송·수신 SMS센터, 단말
필터링 주체	통신사업자	통신사업자	통신사업자	통신사업자, 사용자	사용자
부가 트래픽	발생	발생	미발생	필터링 위치에 따라 결정	필터링 위치에 따라 결정
개인화 서비스	어려움	어려움	-	가능	가능
비고	적용에 따른 발신고객의 동의필요	스팸전송 패턴로그 필요	-	블랙리스트 품질에 성능좌우	학습샘플, 학습과정 필요

- 사용자 개인 성향 반영의 어려움

네트워크 주요 노드인 SMS센터에서 스팸차단을 처리하는 방법[3-6]은 개인 성향 반영이 어렵다. 왜냐하면 스팸 판별 기준은 사람마다 서로 다른데, 하나의 특정 기준으로 모든 사람이 만족할만한 스팸차단 시스템을 구현하는 것은 현실적으로 매우 어렵기 때문이다. 또 사용자의 스팸 판별 기준은 시간에 따라 달라질 수도 있는데 이를 일괄적으로 만족시키기도 어렵다. 예를 들어 어떤 사용자가 임의의 단체에 자진 가입하여 정기적으로 수신하는 메시지를 햄으로 인식하였다고 하자. 그런데 어느 날 이후부터 이 문자 메시지를 스팸으로 분류하여 차단하고 싶다. 이러한 상황을 일괄 기준 시

스템으로는 해결할 수 없다. 따라서 개인 성향을 반영하기 위해서는 사용자마다 서로 다른 스팸 필터를 구현해야 한다.

• 오분류 결과에 대한 업데이트 어려움

요청응답, 송신행태에 의한 차단 기법은 네트워크 노드에서 처리하므로 사용자의 오분류 결과에 대한 업데이트가 어렵다. 블랙리스트 차단 방법은 사용자가 수작업으로 키워드나 블랙리스트 정보를 직접 타이핑해야 하는 번거로움이 수반된다. 또 업데이트하는 키워드에 따라 판별기준이 민감하게 변화하고 양질의 키워드 선정에 따라 성능 결과가 좌우된다.

• 오분류 결과에 대한 책임소재

SMS 문자 메시지는 즉시성과 신뢰성을 전제로 과금되는 통신 수단이므로 오분류 결과에 대한 책임소재가 분명해야 한다. 그러므로 통신사업자가 주체가 되어 네트워크 경로에서 차단하는 것 보다는 사용자 단말기에서 스팸 차단하는 것이 논란의 여지를 줄일 수 있다. 그리고 휴대단말기에서 구현한 스팸필터는 스팸 메시지 확인을 위해 별도의 네트워크 접속이 필요 없다.

• 스팸필터 동작에 따른 부가 트래픽 유발

SMS 센터에서 스팸을 차단하는 방법은 유무선 네트워크에 스팸 필터링을 위한 부가적인 트래픽을 유발시켜 관리 운용 측면에서 비효율적이다. 그리고 개인 정보가 포함된 SMS 정보 트래픽의 유동은 도청이나 해킹에 의한 약점을 제공할 수도 있다.

• 문서분류의 경우 많은 연산량에 대한 부담

자동 문서 분류에 의한 스팸차단은 사용자의 편의성을 제고할 수 있지만, 여러 단계의 전처리과정과 연산량이 많은 학습과정이 필요하다. 그러므로 휴대단말기에서 독립적으로 동작하는 필터링 시스템을 구현하는 것은 매우 부담스러운 일이다. Taufiq[7]는 휴대단말에서 수신 메시지를 바탕으로 지속적으로 학습하고 새로 인입되는 스팸 메시지를 자동으로 차단할 수 있는 시스템을 개발하였다. 하지만 이는 학습 데이터 누적에 따라 특징 단어 개수도 함께 증가하여 '차원의 저주(curse of dimensionality)'[8]에 빠질 수 있다.

조인휘[9]는 SVM(Support Vector Machine) 분류기를 통하여 스팸문자 메시지를 자동분류하고, 특징 벡터를 제한하였지만 PC상에서 동작하는 시뮬레이션 실험이라는 한계가 있었다.

Ⅲ. 관련이론과 제안 시스템

1. 자동 문서분류 시스템

자동 문서분류 시스템은 수많은 텍스트 데이터로 구성되어 있는 문헌 정보들을 자동으로 분류하여 주는 것으로 자동문서 필터링 시스템[10], 이메일 스팸필터[11], 웹 검색엔진, 모바일 검색엔진[12] 시스템 등에 이용되기도 한다.

자동문서 분류 시스템은 패턴인식의 한 응용분야로서 일반적으로 다음과 같은 과정을 거친다.

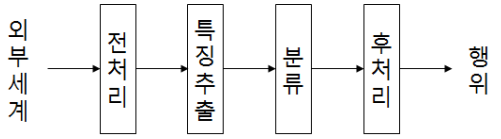


그림 3. 일반적 패턴인식 시스템 처리과정

자동 문서분류 시스템은 텍스트 정보를 입력 받아 분류를 위해 의미 있는 단어(term) 정보를 추출하기 위한 작업을 수행한다. 특징 추출(Feature Extraction or Selection)과정은 분류과정에서 처리할 입력 벡터의 차원을 최소화하면서 변별력 있는 요소를 선정하는 과정이다. 분류과정은 확률 모델, 인공 신경망, SVM, HMM(Hidden Markov Model) 등과 같은 다양한 패턴인식 모델을 적용하여 입력정보에 대한 부류(Classification)를 결정한다. 후처리 과정에서는 부류 결과를 이용하여 시스템의 요구 기능에 연계하여 행위를 취한다.

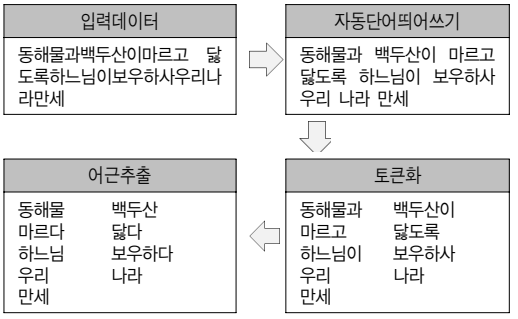
1.1 전처리 단계

전처리 과정은 문헌 정보를 구분할 수 있는 요소 성분 단어를 구분 및 추출하는 과정이다. 그리고 이 과정은 텍스트 데이터를 입력받아 가공 처리하므로 사용언어의 언어학적 특성에 따라, 목적과 환경에 따라 처리 과정을 조정하여야 한다.

보통 입력 문서로부터 단어의 원형을 얻기 위하여 토큰화(tokenization), 어근 추출(steming), 자동 단어 띄어 쓰기, 불용어 제거, 수사어절 정규화 등의 작업을 용도와 환경에 따라 선택적으로 적용한다.

예를 들어 자동 단어띄어쓰기, 토큰화, 어근 추출 처리를 한다면 아래와 같은 결과를 기대할 수 있다.

표 2. 전처리 과정 예시



한글의 경우 위와 같은 전처리를 수행하기 위해서는 어근, 어미, 불용어 등의 사전(dictionary) 데이터베이스가 확보되어 있어야 한다.

1.2 특징 추출 단계

특징 추출 단계는 분류기가 처리해야 하는 특징 벡터 차원을 최소화 하면서 변별력 있는 것으로 선택하기 위한 작업이다. 다시 말하면 분류기 학습 과정에서 문서에 나타나는 여러 단어들 중에 분류를 하는데 중요한 역할을 하는 단어들을 선택하는 작업을 의미한다. 자동 문서 분류 분야에서 특징을 선택하는 하는 방법으로는 문서빈도(Document Frequency), 상호 정보량(Mutual Information), 카이제곱 통계량(Chi Square Statistics)을 사용하기도 한다[13].

문서 빈도는 특징 선택을 하는데 있어서 가장 쉽고 간편한 방법으로 일정 빈도 이상의 문서에 출현한 용어들을 특징으로 추출하는 기법을 말한다. 하지만 한 문서 사이즈가 제한되어 있고 출현 가능한 단어가 많은 환경에서는 정확도를 높이는데 한계가 있다.

상호 정보량은 통계적 언어모형에서 개별 단어간의 연관성을 측정하는데 이용되는 방법으로 두 단어 중 한 단어가 다른 단어에 대해 갖고 있는 정보량을 의미한다. 특정 범주 c에서의 단어 t의 정보량은 다음 식으로 계산된다.

$$I(t,c) = \log \frac{\Pr(t,c)}{\Pr(t) \times \Pr(c)} = \log \Pr(t|c) - \log \Pr(t) \quad (1)$$

상호 정보량 I(t,c)는 t와 c가 서로 독립일 때 0 이 되므로 상호 정보량이 큰 단어순으로 특징을 선택해야 한다. Pr(t|c)가 동일하면 상호 정보량은 Pr(t)에 영향을 크게 받아 희소하게 나타나는 단어에게 높은 값을 주게 되므로 빈도차이가 많은 단어에는 바람직하지 않다.

카이제곱 통계량은 일반적으로 통계분야에서 기대도수와 관측도수의 차이가 유사한지를 판단하는 방법으로 사용되는 것으로 특정 단어와 문서 범주간의 관계도를 측정하는데 사용되며 아래와 같이 표현된다.

$$\chi^2(D,t,c) = \sum_{t \in 0,1} \sum_{c \in 0,1} \times \frac{(N_{t,c} - E_{t,c})^2}{E_{t,c}} \quad (2)$$

N은 문헌 D에 출현할 관찰 빈도이고 E는 평균 빈도이다. 단어 t와 범주 c가 완전히 독립적이면 카이제곱 통계량 역시 0 이 된다. 상호 정보량과 카이제곱 통계량의 차이는 카이제곱 통계량의 값이 정규화된 값이라는 데에 있다. 그러므로 카이제곱 통계량은 단어들 간에 빈도수의 차이가 많더라도 적용할 수 있다.

1.3 분류기

분류기는 양질의 샘플 데이터로 학습과정을 거친 이후, 변경된 분류기의 가중치를 이용하여 새로운 입력 데이터를 분류한다.

분류기로 여러 가지 모델이 사용될 수 있으나 이항 부류 구분에서 가장 좋은 성능을 보이는 모델로는 베이지안 분류기(Naive Bayes)와 SVM 등이 대표적이다.

간단하면서도 좋은 성능을 내는 베이지안 분류기는 확률기반의 모델로서 문서를 이루는 각 단어들이 서로 독립이라는 가정을 전제로 한다. 이 가정 때문에 전체 문서집합에 대한 단어별 빈도수와 문서내의 단어빈도수 정보만 있으면 분류를 할 수 있다. 문서에 대하여 가장 좋은 분류 결과를 돌려주는 베이지안 분류기는 식 (3)과 같다.

$$C_{map} = \arg\max_{c \in C} \hat{P}(c) \prod_{1 \le k \le n_d} \hat{P}(t_k|c) \quad (3)$$

베이지안 모델에서 가장 적합한 범주는 최대 사후 확률(MAP: maximum a posteriori) 범주 C_{map}이다.

SVM은 이진 분류기법으로 두 그룹의 데이터를 구분시키는 분류경계(hyperplane)와 최 근접 학습 데이터와의 거리를 최대화 시키는 최대 여백분류 경계(maximal margin hyperplane)를 찾는 학습 방법이다[8].

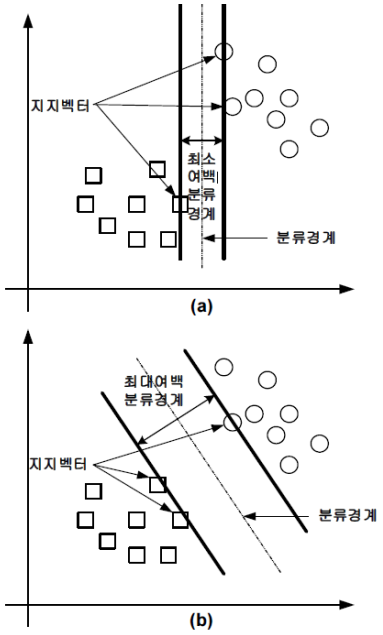


그림 4. SVM을 이용한 데이터 이진분류

분류 경계는 커널 함수로 표현되며 SVM에서 가장 널리 사용되는 커널 함수로는 선형 커널(linear kernel), 다항커널(polynominal kernel), RBF(Radial Basis Function) 커널 등이 있다. 이 중에서 선형 커널은 가장 빠르고 간단한 커널 함수로서 SVM 기반 응용에 많이 사용되고 있다. [그림 4]는 SVM을 이용하여 주어진 학습 데이터를 두 개의 그룹으로 분류한 것이다. [그림 4]에서 (a)는 최소 여백 분류경계에 의한 이진 분류의 예이며, (b)는 최대 여백 분류경계에 의한 이진 분류의 예이다. 이러한 최대(최소) 여백 분류 경계에 가장 근접한 학습 데이터를 지지 벡터(support vector)라고 한다.

2. 제안 시스템

본 연구는 기존 SMS 스팸 차단 서비스 약점을 보완하고 복잡한 연산 과정을 모바일 기기와 사용자 PC(Personal Computer)간에 분산하여 동작하는 메시지 내용 학습기반의 스팸 필터링 모델을 제안한다.

2.1 제안 시스템의 구성

본 연구에서는 스마트폰 사용자가 개인 컴퓨터에 정기적으로 동기화 작업을 수행하는 것에 착안하여 스팸 차단 시스템을 디자인하였다. 스마트폰은 동기화를 통하여 주요 정보를 PC에 백업하거나, 어플리케이션의 주요 파라미터를 갱신하거나 배터리를 충전하기도 한다. 제안 시스템에서 많은 연산이 필요한 분류기의 학습 과정은 모두 사용자 컴퓨터에서 처리하고, 스마트폰은 필터링 연산 과정만을 수행 한다. 사용자의 컴퓨터는 학습 데이터를 전달 받아 분류기의 결정 평면 가중치 값과, 키워드 등의 정보를 스마트폰으로 전달한다. 스마트폰은 새로 유입되는 SMS에 대해서 간단한 전처리 과정으로 생성된 특징벡터를 인식 처리하여 스팸을 구별한다.

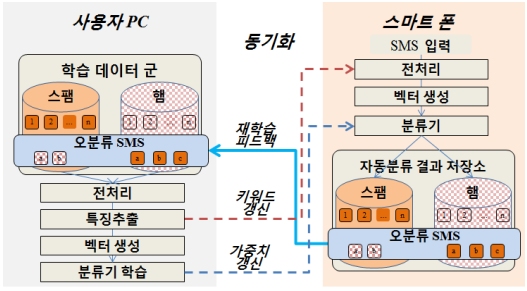


그림 5. 제안 시스템 구성도

2.2 제안 시스템의 특징

제안하는 SMS 스팸 필터링 시스템은 고비용의 연산 과정을 사용자 인증 컴퓨터와 스마트폰 사이에 분산시킴으로써 다음과 같은 특징이 있다.

- 사용자 성향 반영
  - 스마트폰에서 수신메시지 데이터를 기반으로 학습
  - 오분류 결과에 대하여 사용자가 확인하고 학습데이터에 반영하여 결정평면 가중치를 갱신

- 일정 자원 소모
  - 학습 데이터가 증가하더라도 스마트폰이 처리해야 하는 정보의 종류 및 크기는 일정
  - 차원의 저주를 방지하기 위해 사용자 PC에서 특징 벡터를 선정함

- 부가 트래픽 미발생
  - 스팸차단 실행 및 스팸메시지 박스 확인과 관련하여 부가적인 트래픽을 네트워크에 발생시키지 않음

Ⅳ. 설계 및 구현

1. 동기화 교환 정보

사용자 컴퓨터와 스마트폰 사이에 주고 받는 정보는 아래 [표 3]과 같이 스마트폰이 PC에 전달하는 것과 PC가 스마트폰에 전달하는 것으로 구분된다.

표 3. 동기화 교환 정보 목록

구분	교환 정보 목록
스마트폰 -> PC	·스마트폰에서 수신된 오분류 메시지
PC-> 스마트폰	·특징 벡터를 구성하는 특징 키워드 리스트 ·학습 완료된 분류기 가중치

스마트폰이 사용자의 컴퓨터로 전달하는 정보는 오분류된(misclassified) 문자 메시지이다. 스마트폰의 스팸 필터 결과에 대해 사용자가 오류판정을 내린 SMS 데이터를 동기화 시점에 사용자 컴퓨터로 전달한다. 이때 스마트폰은 컴퓨터로 전송할 오분류 문자 데이터에 라벨(label)을 다시 달아 분류기의 학습 샘플 집합에 반영한다.

사용자 컴퓨터는 스마트폰으로부터 전달 받은 피드백 데이터에 대해 추가적인 학습을 마친 후, 특징추출 키워드 리스트, 분류기의 가중치 등의 두 가지 정보를 스마트폰으로 전달한다. 특징추출 단어는 전처리 과정을 거친 단어들인데 이는 스팸, 햄을 결정하는 데 주요한 역할을 하는 단어들을 일정한 기준에 의해 선정한 것이다. 이 단어 개개는 분류기에게 한 차원이 되므로, 차원의 저주에서 벗어나기 위해서 가능한 특징추출 단어를 줄여야 한다. 분류기의 가중치는 일종의 벡터 테이블로서 Naïve Bayes 모델에서는 각 단어들의 출현 빈도에 근거한 확률 값, 선형 SVM에서는 결정 평면에 대한 가중치이다.

2. 전처리 및 특징 추출

본 논문에서는 스마트폰에서 이루어지는 한글 문자 메시지의 전처리 과정의 부담을 최소화하기 위하여 분산처리 방법을 사용하였다. 사용자 컴퓨터는 한국어 형태소 분석기(KLT[1], Korean Language Technology)를 이용하여 주요 색인어를 검출하고, 이 색인어 중에서 특징 벡터로 선정된 단어들만 스마트폰으로 전달한다. 스마트폰은 전달받은 색인어를 이용하여 새로 입력되는 SMS에 대하여 아주 간단한 키워드 매칭 방식으로 특징 벡터를 검출한다.

1 KLT http:// nlp.kookmin.ac.kr/HAM/kor

표 4. 세부 전처리 과정

세부 전처리 과정	PC	Phone
특수문자 제거	✓	✓
공백단위로 분리	✓	✓
어근추출	✓	
자동 띄어쓰기	✓	
키워드 매칭		✓
숫자어 인식	✓	✓
특징 추출	✓	

전처리 과정은 텍스트 문장 형태의 SMS를 특징 벡터 추출이 용이하도록 가공하는 것이다. 이 단계는 특수 문자 제거, 토큰 분리(tokenization), 어근 추출(stemming), 색인단어 검출, 숫자어 구분 등의 과정을 거친다. 유니코드 기반의 한국어의 경우, 160byte 메시지 안에 포함된 단어의 수가 영어에 비해 매우 적고, 토큰화 과정만으로 색인어를 분리할 수 없다. 따라서 조사, 어말어미 등을 분리하는 어근추출(steming) 과정을 거쳐야 하는데, 이 과정은 스마트폰으로 처리하기에는 너무 복잡하므로 [표 4]와 같이 그 역할을 PC가 대신하였다.

특수 문자와 한 글자 단어는 스팸과 햄을 구분하는데 큰 역할을 하지 않고 오히려 색인어 변형의 수단으로 사용될 수 있으므로 제거한다. 하지만, 다음과 같은 심볼 및 변형 문자는 문자 메시지에 유용하게 사용하는 이모티콘이므로 단어 추출시 참조 사전으로 사용한다. 그리고 스팸 메시지에서만 등장하는 특정 키워드들은 아래 [표 5]과 같이 참조 사전으로 이용하여 한 가지 단어로 인식하도록 한다. 이런 과정은 스팸 키워드의 등장을 집중시켜 분류 정확도를 향상 시킬 수 있다.

표 5. 전처리 과정에서 사용한 참조 사전

구분	예시
이모티콘	^^ ^.^ -_.- -.- (_ _) (- -)
스팸 키워드	060 080 2야기 B아 capital C알 가승인 고래 긴급자금 ㄷㅐㄹㅣ운전 ㄷㅐ리운전 ㄷㅐ출 대ㄹㅣ운전 대리운전 대출 대출 란제리 물롤렛 릴게임 무료상담 무료충 무이자 밀봉 ㅂrㄷr 바다이야기 바카라 바카라 바카렛 박카라 배팅 블랙잭 블랙잭 블랙짹 신속입금 씨알 ㅇㅣㅈㅏ ㅇㅣ자 야마토 이자 입금대기 최저금리 최하이율 ㅋㅐㅍ탈 ㅋㅔㅍ털 ㅋㅔㅍㅣ탈 카지ㅣ노 카지노 카지노 캐피탈 캐피탈 케피탈 케피털 풀싸롱 풀코스 풀클럽 하드풀

토큰 분리 과정에서는 SMS 문장을 공백을 기준으로 나누어 하나의 단위로 구분한다. 토큰은 문법적으로 더 이상 나눌 수 없는 기본적인 언어 요소를 말하는데, 한글의 경우에는 체언과 조사를, 본용언과 보조용언을 분리하는 과정을 거쳐야 한다. 그리고 한글 스팸 메시지는 의도적으로 띄어쓰기 문법을 지키지 않고 스팸 키워드 사이에 특수문자를 끼워 넣는 경우가 많다. 따라서 색인 단어 추출을 위해 특수 문자, 공백을 모두 제거한 가공된 문장을 KLT를 이용하여 형태소 분리를 실시하였다.

숫자어 구분 단계에서는 숫자어, 숫자를 가독하여 돈과 연관된 단어, 숫자어, 유사 숫자어 3가지로 구분하였다[14]. 스팸의 대부분이 대출, 사행성 게임 등으로 돈, 숫자와 관한 내용이 많기 때문이다.

3. 벡터 생성

특징 추출 과정은 분류기에 입력되는 특징 벡터의 차원수를 줄이기 위해 스팸, 햄 구별에 중요한 역할을 하는 단어만을 선정하는 과정이고 PC에서만 수행된다. 특징 추출 없이 모든 색인어를 벡터 요소로 이용하면 차원의 저주로부터 자유롭지 못하여 너무 많은 컴퓨팅 자원을 소모해야 한다. 본 논문에서는 카이 제곱 통계량을 이용하여 특징 단어를 선정하였다. 특징 단어 선정을 위해 각 단어에 대해 식 (2)로 용어를 순위화하고 값이 큰 단어를 선정하였다.

4. 분류기

SMS 스팸 차단은 이메일 스팸차단 기술과 매우 유사한 기법들을 적용할 수 있다[6]. 본 논문에서는 이메일 스팸차단에 가장 훌륭한 성능을 보이는 Naïve Bayes과 선형(linear) SVM을 분류기로 적용하였다. Naïve Bayes 알고리즘으로 각 단어에 대해서 사후 확률을 구하여 스팸, 햄 판정을 한다.

즉, Naïve Bayes 확률 모델은 식 (3,4)에 결과에 따라, 판별 대상 문자메시지를 가장 큰 값을 보이는 군(Class)으로 간주한다.

$$\text{if } \hat{P}(C_{spam}|d_{SMS}) > \hat{P}(C_{ham}|d_{SMS}) : spam$$
$$\text{if } \hat{P}(C_{spam}|d_{SMS}) \leq \hat{P}(C_{ham}|d_{SMS}) : ham \quad (4)$$

SVM은 스팸, 햄을 구분하는 결정 초평면을 구하는데 결정적인 역할을 하는 경계 요소들 사이의 마진(margin)을 최대한 확보하는 알고리즘이다. SVM은 커널 함수의 종류에 따라서 선형, 비선형으로 구분되는데, 비선형 SVM은 차원의 수가 증가할수록, 서포트 벡터 요소가 많을수록 연산량이 증가하는 단점이 있다. 게다가 수백 차원 정도의 시스템일 경우에는 비선형 SVM보다 선형 SVM이 보다 더 좋은 정확도를 보이기도 한다. 따라서 본 연구에서는 선형 SVM을 사용하고, SVM-JAVA[2]이라는 공개 프로그램을 사용하였다. SVM-JAVA는 John C. Platt의 SMO(Sequential Minimal Optimization)를 자바로 구현한 것이다.

V. 실험 및 고찰

1. 데이터 수집

본 연구의 정확한 실험을 위하여 가능한 많은 수의 한글 문자 메시지를 수집해야 했다. 스팸 문자 메시지는 무선 네트워크 서비스 사업자인 KT가 제공하는 스팸 필터링 시스템[3]을 통하여 자동으로 차단되는 스팸 메시지를 22명의 자원자로부터 1,055개를 제공 받았다. KT가 제공하는 스팸 차단 시스템은 사용자가 개별적으로 입력한 키워드 또는 [표 6]과 같은 스팸 키워드를 이용하여 키워드 매칭 방식으로 동작한다. 자발적으로 제공받은 스팸 메시지는 약 3개월 동안 누적된 데이터로서 22명의 사용자가 정확히 스팸으로 확인한 것이다. 햄 메시지는 스팸과 달리 수집하기가 어렵다. 스팸 메시지는 사생활 정보가 거의 포함되어 있지 않아서 자발적 데이터 제공자를 구하기가 용이하였지만, 햄 메시지는 공개하기 어려운 사적 대화 내용을 비롯하여 개인정보가 포함되어 있어서 수집하기가 불가능했다. 따라서 햄 메시지를 대체할 1,493개의 정보를 트위터 웹사이트에서 수집하였다. 트위터는 익명의 사용자가 대화체 형태의 글과 여러 가지 특수문자를 이용한 이모티콘을 사용하여 실제의 문자 메시지와 매우 유사하다.

표 6. 숫자어 종류와 예시

숫자어 종류	예시
돈과 연관된 단어	백만원, 100만원, 1,000,000원
숫자어	백만,100만, 1,000,000
유사 숫자어	9천8만원, 구천팔만

2. 실험 및 결과

2.1 일괄처리 테스트

이 실험은 다른 연구자들이 실시한 방법과 유사하다. 2,548개의 데이터 샘플을 [표 7]와 같이 훈련 집합군과 테스트 집합군으로 6대 4의 비율로 나누고, 일괄적으로 훈련, 테스트를 실시하였다. 1,528개의 데이터를 전처리 후 6,565개의 특징 단어가 생성됨을 확인할 수 있었다.

표 7. 실험 데이터 구분

구분	햄 메시지	스팸 메시지	합 (비중)
훈련 집합	895	633	1,528 (60%)
테스트 집합	598	422	1,020 (40%)
총 계	1,493	1,055	2,548 (100%)

카이 제곱 통계량 기법을 이용하여 선택되는 특징 단어 개수를 변화시켜 가면서 정확도와 훈련시간의 변화를 관찰하였다. 여기서 훈련시간은 특징 벡터를 통한 분류기의 훈련 시간과 문자 메시지의 전처리 과정을 포함한 시간이다. 전반적 인식 결과는 [그림 6]과 같이 선형 SVM이 평균 95.3%, Naive Bayes는 평균 91.1%를 보여 선형 SVM이 다소 우수하였다. 차원의 수가 지속적으로 증가할 때 Naive Bayes나 SVM 두 방법 모두 인식률이 아주 완만하게 증가함을 확인할 수 있었다.

두 모델의 훈련 시간 측면에서는 [그림 7]에서 확인할 수 있듯이 SVM의 경우 차원 수가 증가할수록 학습시간이 지수 형태로 증가함을 확인할 수 있었다. 당 실험 결과에 따르면 인식 결과가 좋은 SVM을 이용하려

2 http://iis.hwanjoyu.org/svm-java
3 http://mobile.olleh.com/index.asp

면 훈련 집합의 개수에 따른 적당한 특징 벡터 차원의 수를 조정해야 함을 알 수 있었다. 당 실험 데이터에 적당한 차원의 수는 1,000을 넘지 않을 때 양호한 결과를 보임을 결과 그래프를 통해서 알 수 있었다.

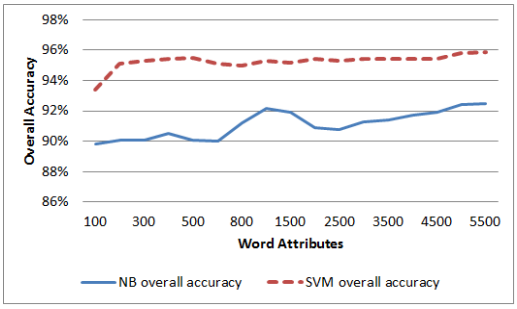


그림 6. 특징 벡터수에 따른 인식률 추이

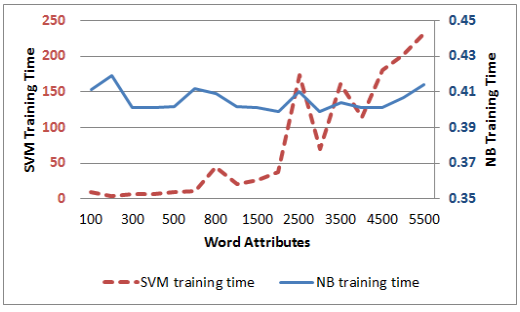


그림 7. 특징 벡터수 변화에 따른 훈련시간 추이

2.2 오분류 메시지 업데이트

2.1의 일괄처리 테스트는 실제상황에서 발생하기 어려운 조건이다. 왜냐하면 1천여 개가 넘는 문자 메시지를 한 사람이 스마트 폰에 저장하고 다니는 사람은 현실적으로 없기 때문이다. 따라서 두 번째 실험에서는 [그림 5]와 같이 본 연구의 제안 모델을 최대한 유사하게 시뮬레이션하기 위한 실험 조건을 조성하였다. 모바일 폰에 50개의 스팸과 50개의 햄 메시지가 이미 저장되어 있는 것으로 가정하고 이 100개의 데이터로 분류기를 학습시켰다. 이후 50개씩 새로운 메시지를 학습된 분류기를 통과하여 스팸 또는 햄으로 구분하였다. 이때 스마트폰에서는 오분류된 데이터를 따로 저장하고 있다가 동기화 시점에 그 데이터를 사용자 컴퓨터에 전달하였다. 동기화 프로세스는 50개의 메시지를 처리한 이후에 정기적으로 실시하고 컴퓨터는 전달받은 메시지를 이용하여 재학습 이후 스마트폰으로 분류기 가중치 정보, 특징 단어 키워드 등을 업데이트 하였다.

이 실험 결과 Naive Bayes 모델을 이용한 인식률은 첫 번째 실험에 비교할 때 상당히 좋은 변화가 있었다. 인식률 결과가 항상 일정하진 않았지만 아래 [그림 8]과 같이 매번 90%이상의 결과를 보였고, 평균 95.3%의 인식 결과를 얻을 수 있었다.

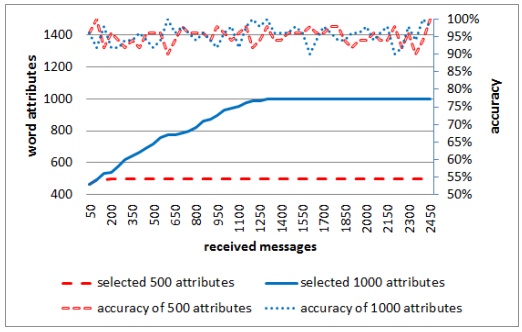


그림 8. Naive Bayes의 피드백 업데이트 결과

Naive Bayes를 이용한 당 실험에 이용된 분류기의 가중치 정보, 특징 키워드 정보의 크기는 9 Kbyte 내외로 일정하였고, 스마트 폰에서 메시지 한 개를 처리하는 데 걸리는 시간은 0.08초가 걸렸다.

SVM의 경우에는 평균 95.4% 정확도로 첫 번째 실험에 비해 큰 개선이 없었다. 하지만 오분류 메시지만을 학습 샘플 데이터로 업데이트하여 분류기의 학습 소요시간은 크게 개선되었다. 분류기의 차원 수가 300일 경우에는 학습 시간이 평균 0.8초가 걸렸고, 차원수가 150일 경우에는 평균 0.5초가 걸렸다. 그리고 SVM 가중치와 특징 키워드 단어의 총 크기는 3.8 Kbyte로서 스마트폰에서 한 개의 메시지를 처리하는 시간은 0.21초가 소요되었다.

[그림 9]에서 SVM의 인식 정확도는 동기화 재학습 이후 급격하게 향상되었다. 그리고 150 차원이나 300 차원 모두 300개 문자 메시지를 받은 시점까지 인식 정확도가 같음을 확인 할 수 있다. 이는 특징 단어가 특징

벡터 150에 다다를 때까지 모두 같기 때문이다. 특징 단어 개수가 최대에 다다른 이후에는 각 단어들은 특징 선정을 위한 과정을 각각 거치기 때문에 인식률의 차이가 나타난다. 아래 그림과 평균치 결과에 의하면 300차원의 SVM 시스템이 좀 더 양호한 결과를 보였으나, 300 차원의 시스템이 150 차원의 시스템 보다 더 좋은 시스템이라고 할 수는 없다. SVM 시스템의 차원의 수는 학습 샘플 데이터의 개수, 하드웨어 사양 등의 관점에서 최적의 조건을 선정해야 한다.

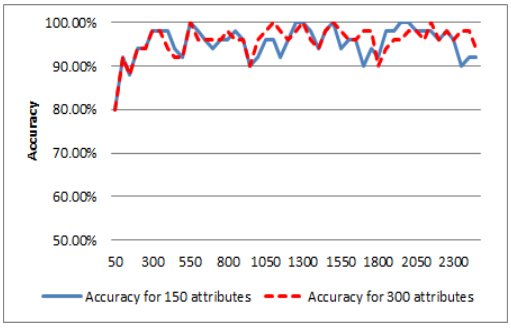


그림 9. SVM의 피드백 업데이트 결과

2.3 기존 연구 비교

SMS 스팸 필터 분야에서 기존 연구와 비교하는 일은 쉬운 일이 아니다. 영상 연구 분야는 샘플 데이터를 서로 공유 하여 각종 실험 결과를 기존 연구와 비교하는 것이 가능한데 비해, 당 분야는 데이터 샘플을 공유하는 일이 드물고, 사용 언어에 따라, 지역에 따라 스팸 메시지의 특성이 달라진다. 게다가 언어가 다르면 특징 단어를 추출하는 언어학적 성격이 완전히 달라지므로 [13] 모든 상황에 적용할 수 있는 범용스팸 필터를 개발하는 것은 불가능하다. 본 논문에서는 제안 시스템과 유사한 연구 모델을 선정하고 동일한 한국어 데이터 샘플을 이용하여 동일 조건에서 실험하고 인식 정확도를 비교하여 보았다.

[표 8]은 제안 시스템이 어떠한 점에서 기존 연구와 서로 다르고 유사한 것인지를 보여준다.

표 8. 기존 연구와 비교

구분	Taufiq[7]	조인휘[9]	제안 시스템
분류기	Naive Bayes	SVM	Naive Bayes, SVM
전처리 수행	Mobile 기기	PC	PC, Mobile기기
특징추출/수행	–	카이제곱통계량/ PC	카이제곱통계량 /PC
분류 수행	Mobile 기기	PC 시뮬레이션	Mobile 기기
사용언어/부호	알파벳/ ASCII 코드	한글/ 유니코드	한글/ 유니코드
한글스팸	처리불가	처리가능	처리가능
휴대기기 동작	가능	현실적 불가	가능

Taufiq[7]은 알파벳 영문 스팸을 필터링 할 수 있으나, 한글 스팸은 처리가 불가능하다. 이 방식은 전처리 과정으로 토큰화 과정만을 수행하는데, 똑같은 방식을 한글 스팸 필터링에 적용할 경우 인식율이 60%이하로 현저하게 낮아지게 된다. 왜냐하면 한글은 영문과 달리 체언에 조사가 결합되어 동일한 어절의 출현 빈도가 매우 낮기 때문이다.

조인휘[9]는 카이제곱통계량으로 선정된 특징벡터를 SVM에 적용하였으나 휴대기기에 어떻게 적용할 것인지에 대한 내용은 부재하다. 이 방식의 모든 과정을 모바일 기기에서 독립적으로 수행하는 것은 현실적으로 불가능하다. 스마트폰의 CPU, 메모리, 배터리 성능이 데스크탑 수준으로 향상된다면 가능하겠지만, 현존하는 스마트폰이 처리하려면 매우 많은 연산시간과 배터리 소모를 감수해야 한다.

조인휘[9] 방식의 모든 단계를 스마트 폰에서 수행한다고 가정하였을 때, 그 소요시간이 현실적으로 가용한 정도인가를 관찰하기 위하여 특징 선정과 SVM 학습에 소요되는 시간을 측정하는 실험을 실시하였다. 여러 단계 중 카이제곱 통계량을 통한 특징 선정과 SVM 학습과정이 가장 많은 연산이 소요되므로 경과시간 측정을 통해 가용수준 정도를 판단할 수 있다.

표 9. 특징선정과 SVM 학습과정의 소요시간

특징 선정 + SVM 학습	PC	Mobile
1308차원 중 150차원 선정후 SVM학습	4.85 초	53.3 초
1308차원 중 200차원 선정후 SVM학습	5.15 초	98.8 초
1308차원 중 250차원 선정후 SVM학습	3.54 초	44.7 초
1308차원 중 300차원 선정후 SVM학습	4.45 초	83.6 초

위의 결과는 햄 150개, 스팸 100개로 구성된 1,308개 단어(차원)의 이진 벡터 정보에서 150~300개 단어(차원)를 선정하고 SVM 학습 수렴에 소요된 시간을 측정한 것이다. PC에서는 4~5초 내외의 시간이 소요되었지만, 스마트 폰에서는 50~100초 내외의 시간이 걸렸다. 모바일 기기의 제한적 배터리 사용, 프로그램 응답속도 등을 감안할 때, 이 결과는 모든 과정을 모바일 기기에서 처리하는 것이 현실적으로 어렵다는 것을 보여준다.

위 [표 8]의 세 모델의 인식률을 비교하기 위하여 동일한 PC에서 동일한 샘플 데이터를 이용하여 실험을 진행하였다. Taufiq[7]의 모델은 특징 추출 과정 없이 토큰화 처리만 수행하였고, 조인휘[9]의 모델은 오분류 데이터에 대한 피드백 없이 학습, 분류 과정을 수행하였다. 당 연구에서 수집한 데이터 샘플로 실험하여 [표 10]과 같은 결과를 얻을 수 있었다.

표 10. 인식률 비교

Accuracy	NB	SVM
Taufiq[7]	91%	-
조인휘[9]	-	92.7%
제안 시스템	95.3%	95.4%

3. 실험 환경

본 논문의 모든 실험은 다음과 같은 환경에서 실행하였다. 컴퓨터와 스마트폰의 동기화 과정은 관련 정보를 포함하고 있는 파일을 옮기는 것으로 시뮬레이션 하였다.

표 11. 실험 환경

구분	하드웨어 및 운영체제
스마트 폰	Qualcomm® QSD8250™, 1 GHz Processor Android™ 2.1 (clair) Operating System ROM Memory 512MB and RAM 512MB 4.MicroSD™ memory card (SD 2.0 compatible)
컴퓨터	Intel core2 Duo CPU 2.93 Ghz, RAM 3.0 GB Windows 7, Java SE 1.6

VI. 결론 및 향후과제

본 논문에서는 스마트폰의 자원을 가능한 적게 사용하면서 사용자의 취향을 반영할 수 있는 스팸 문자 메시지 차단 시스템을 제안하였다. 제안 시스템은 분류기로 Naive Bayes, SVM을 적용하여 양호한 수준의 인식 결과를 보였다. 기존 연구와 차별적으로 학습과 인식 과정을 분산 처리하여 사용자 선호도 반영, 인식 성능 향상, 일정한 자원 사용이라는 성과를 얻을 수 있었다. 본 논문의 제안 모델은 적은 수의 임의의 학습 데이터로 시작하더라도 시간이 경과하고 몇 번의 동기화 업데이트 이후에 사용자의 스팸 기준 취향에 빠르게 적응할 수 있다. 이에 더하여 본 제안의 동기화 프로세스는 인증된 사용자 컴퓨터와 스마트폰 사이만 이루어지므로 개인 정보 유출 등의 문제로부터 자유롭다.

본 논문에서는 PC와의 동기화를 통한 스팸 분류기 학습 정보를 교환하였으나, 향후 PC 대신 클라우드 컴퓨팅을 통한 분산 처리도 가능할 것으로 보인다. 그리고 스팸 판별시 스마트폰에 저장되어 있는 전화번호부 정보 존재 유무 과정을 한 번 더 거친다면 인식 정확도는 더 향상 될 수 있을 것이다.

참 고 문 헌

[1] http://www.smartwork.go.kr.

[2] 방송통신위원회, *스팸방지 종합대책*, 2011, http://spam.kisa.or.kr/kor/notice/noticeView.jsp?mode=view&p_No=10&b_No=10&d_No=64.

[3] P. He, Y. Sun, and W. Zheng, "Filtering Short Message Spam of Group Sending Using CAPTCHA," in Proc. of Workshop on Knowledge Discovery and Data Mining, Adelaide, Australia, pp.558-561, 2008.

[4] X. Hu and F. Yan, "Sampling of Mass SMS Filtering Algorithm Based on Frequent Time-Domain Area," in Proc. of Third International Conference on Knowledge Discovery and Data Mining, Phuket, Thailand, pp.548-551, 2010.

[5] W. Qian, H. Xue, and W. Xiayou, "Studying of

Classifying Junk Messages Based on The Data Mining," in Proc. of International Conference on Management and Service Science, Beijing, China, pp.1-4, 2009.

[6] K. Yadav, P. Kumaraguru, A. Goyal, A. Gupta, and V. Naik, "SMSAssassin: Crowdsourcing Driven Mobile-based System for SMS Spam Filtering," in Proc. of 12th Workshop on Mobile Computing Systems and Applications, 2011.

[7] M. Taufiq, "Independent and Personal SMS Spam Filtering," in Proc. Of 11th IEEE International Conference on Computer and Information Technology, Sep. 2011.

[8] C. M. Bishop, Pattern Recognition and Machine Learning, Springer-Verlag, 2006.

[9] 조인휘, "휴대폰 SMS를 위한 SVM 기반의 스팸 필터링 시스템", 한국통신학회논문지, 제34권, 제9호, pp.908-913, 2009.

[10] 손기준, 임수연,"베이지안 분류기를 이용한 문서 필터링", 한국콘텐츠학회논문지, 제5권, 제3호, pp.227-235, 2005.

[11] 임양원, 임한규,"사용자 패턴을 이용한 지능형 e-메일 시스템의 연구", 한국콘텐츠학회논문지, 제6권, 제1호, pp.65-72, 2006.

[12] 조종근, 하상은, "모바일 환경에서 파일 검색 엔진을 위한 효과적인 방식", 한국콘텐츠학회논문지, 제8권, 제11호, pp.41-48, 2008.

[13] C. D. Manning, *Introduction to Information Retrieval*, Cambridge University Press, 2009.

[14] 강승식, "한국어 수사어절의 유형 분류 및 정규화", 한국정보과학회 1999년도 가을 학술발표논문집, 제26권, 제2호, pp.187-189, 1999.

저 자 소 개

이 승 재(Seung-Jae Lee) 정회원

- 1995년 8월 : 한국전기통신공사
- 1996년 2월 : 전남대학교 전산학과(학사)
- 2010년 3월 ~ 현재 : 전남대학교 전자컴퓨터공학과(석사과정)
- 1995년 8월 ~ 현재 : KT 글로벌 기업고객부문(차장)

<관심분야> : 유무선 네트워크 관리, 자연언어 처리

최 덕 재(Deok-Jai Choi) 정회원

- 1982년 2월 : 서울대학교 컴퓨터공학과(학사)
- 1984년 2월 : KAIST 전산학과(석사)
- 1993년 ~ 1995년 : University of Missouri-Kansas City Computer Science and Telecommunication Program(박사)
- 1996년 ~ 현재 : 전남대학교 전자컴퓨터공학과(교수)

<관심분야> : 상황인식, Pervasive Computing, Future Internet, Sensor Network, IPv6