

동시출현 단어분석 기반 스팸 문자 탐지 기법*

오 하 영^{†*}
송실대학교

Coward Analysis based Spam SMS Detection Scheme*

Hayoung Oh^{†*}
Soongsil University

요 약

스팸 데이터 셋은 통상적으로 공개적으로 구하기 어렵고 기존 연구들은 대부분 스팸 이메일에 초점이 맞춰져 왔기 때문에 스팸 문자 메시지 자체 특성을 분석하는데 한계가 있었다. 스팸 이메일 특성 분석 활용 및 데이터 마이닝 기법 등의 활용을 통한 기존 연구들이 있었지만, 영향력이 높은 단일 단어를 활용한 스팸 문자 탐지 기법에 한정되어 있다는 한계점이 있다. 본 논문에서는 싱가포르 대학교에서 공개적으로 공개한 스팸 문자메시지를 다 각도에서 실험 및 분석하여 스팸 문자의 특성을 밝히고 동시출현 단어분석 기반의 스팸 문자 탐지 기법을 제안한다. 성능평가 결과, 제안하는 기법의 거짓 양성률과 거짓 음성률이 2%미만임을 보였다.

ABSTRACT

Analyzing characteristics of spam text messages had limitations since spam datasets are typically difficult to obtain publicly and previous studies focused on spam email. Although existing studies, such as through the use of spam e-mail characterization and utilization of data mining techniques, there are limitations that influence is limited to high spam detection techniques using a single word character. In this paper, we reveal the characteristics of the spam SMS based on experiment and analysis from different perspectives and propose coward analysis based spam SMS detection scheme with a publicly disclosed spam SMS from the University of Singapore. With the extensive performance evaluations, we show false positive and false negative of the proposed method is less than 2%.

Keywords: Properties of spam SMS, coward analysis, false positive, false negative

1. 서 론

스팸머는 이메일 혹은 문자 등을 활용해서 원하는 목적을 악의적으로 달성한다. 초창기 인터넷의 보급은 스팸 이메일의 문제를 주로 야기시켰지만 최근 스마트폰의 사용량의 증가는 스팸 이메일은 물론 스팸

문자 메시지 양의 급격하게 증가시켰다 [1][2]. 스팸 문자 데이터 셋은 공개적으로 구하기 어렵고 구한다고 해도 스팸인지 정상인지 구별 (labeling)이 확실하지 않기 때문에 스팸 문자 메시지 특성 분석활용에 한계가 있었다.

스팸 문자에 관한 연구는 아래와 같이 크게 3분야로 나눌 수 있다. 첫째, 스팸 이메일의 연구 결과를 스팸 문자 연구로 연결시킨 기존 연구들이다 [3][4]. [3]의 저자들은 스팸 이메일 탐지를 위한 베이지언 분류 모델을 확장해서 스팸 문자에 적용했다. [4]의 저자들은 스팸 이메일에 활용된 특징(features)들을 활용하여 스팸 문자 탐지에 활용하

Received(09. 24. 2015), Modified(1st: 01. 18. 2016, 2nd: 05. 19. 2016), Accepted(06. 08. 2016)

* 이 논문은 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 한국연구재단에서 부여한 과제번호: NRF-2014R1A1A1003562)

† 주저자, hyoh@ssu.ac.kr

‡ 교신저자, hyoh@ssu.ac.kr(Corresponding author)

려고 했다. 하지만, 스팸 이메일과 달리 문자 메시지는 최대 160문자만 허용되고 이메일에서 주로 사용된 특징 단어와 문자에서 사용되는 특징 단어는 성격이 다르기 때문에 스팸 이메일 탐지에 적용됐던 분류 기법을 그대로 스팸 문자 탐지에 활용하는 것은 성능에 한계가 있다. 둘째, 용어 기반 스팸 문자 탐지 기법 (content based spam SMS detection scheme)들이다 [5][6]. [5]의 저자들은 거짓 양성을 일으키는 용어들에게 부정적인 가중치를 적용 후 비용 함수 (cost function)을 다시 계산하여 기본적인 베이저언 분류 기법을 확대했다. [6]의 저자들은 스팸 문자 메시지에 등장한 단어들의 빈도수를 반영한 첨자 모델 (index model)을 정의하고 재귀함수로 모델링하여 스팸 문자 탐지에 활용했다. 하지만, 단일 단어의 출현 빈도수만 고려했기 때문에 스팸 문자 메시지 탐지율에 한계가 있다. 셋째, 기계학습 기반 스팸 문자 탐지 기법에 관한 연구들이다 [7][8]. [7]의 저자들은 스팸 문자 탐지에 실시간성을 반영하기 위해서 시간 영역 (time-domain)에서 각 단어의 사용 빈도수를 분석 후 특정 빈도수 이상 출현하는 단어들만 고려해서 시간을 단축했다. [8]의 저자들은 처음으로 일대일 사용자들이 주고받는 문자 메시지를 분석하여 방향성 그래프 (directed graph)에 표현 후 장기적으로 보내는 사용자의 패턴을 오프라인으로 분석했다. 또한, 온라인으로 메시지를 주고받는 사용자들의 패턴을 동시에 분석해서 스팸 문자 탐지에 활용했다. 마지막으로, [9]의 저자들은 스팸 문자 분류 기법에 관한 전체적인 소개와 데이터 셋을 구성하는 방법 및 현재 연구의 흐름들을 정리하며 다양한 스팸 탐지 기법을 분류하고 있다. 하지만, 위의 모든 방식들은 스팸 문자 메시지를 분류하기 위해서 특정 단어 분류 기법, 특정 단어의 사용 빈도수, 가중치 적용 등에 초점을 맞추고 두 개 이상의 동시출현 단어분석을 활용하지 못했다.

본 논문은 싱가포르 대학에서 공개적으로 공개한 스팸 문자 메시지를 충분히 실험하여 스팸 문자 메시지의 특성을 분석하고, 스팸 문자 탐지의 정확도를 높이기 위해 동시출현 단어분석을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서 실험 데이터 셋 소개를, 3장에서는 실험 기반 스팸과 정상 문자 메시지의 차이점을 소개한다. 4장에서는 본 논문에서 제안하는 동시출현 단어분석기반 스팸 문자 탐지 기법에 대하여 설명한다. 5장에서는 성능평가 결과를 보여주며, 마지막으로 6장에서는 본 논문을

정리하며 마친다.

II. 실험 데이터 셋

본 논문에서는 싱가포르 대학 (National University of Singapore: NUS)에서 공개적으로 제공한 Spam과 Ham으로 구별 (labeling)된 문자 메시지 데이터 셋 (SMS Spam Collection)을 활용 한다 [3][10]. Fig. 1은 SMS Spam Collection 데이터 셋의 구성과 예시를 보여준다. 직관적이지만 Ham문장들은 Spam문장들에 비해서 대략 6.4배 정도 많음을 알 수 있다.

Labeling	Amount
# of spam sentences	747
# of ham sentences	4,827
# of total sentences	5,574
# of common meaningful words in spam and ham sentences	405
# of meaningful words in spam sentences	327
# of total words in spam sentences	734
# of meaningful words in ham sentences	2313
# of total words in ham sentences	2718

```

1 ham Go until jurong point, crazy.. Available only in bugis
2 ham Ok lar... Joking wif u oni...
3 spam Free entry in 2 a wily comp to win FA Cup final tkts 21st
4 ham U dun say so early hor... U c already then say...
5 ham Nah I don't think he goes to usf, he lives around here
6 spam FreeMsg Hey there darling it's been 3 week's now and no word
7 ham Even my brother is not like to speak with me. They treat me
:
5569 spam REMINDER FROM 02: To get 2.50 pounds free call credit and
5570 spam This is the 2nd time we have tried 2 contact u. U have won
5571 ham Will u b going to espianade fr home?
5572 ham Eity, * was in mood for that. So...any other suggestions?
5573 ham The guy did some bitching but I acted like i'd be interested
5574 ham Rotfl. Its true to its name

```

Fig. 1. Data set of SMS Spam Collection from NUS)

III. 실험 기반 스팸과 정상 문자 메시지의 특성 비교

3.1 기초 실험 결과

spam과 ham의 특성 및 차이점을 비교해보기 위해 아래와 같이 다양한 간단 실험들을 수행했다.

- 단어 개수
- 문장 길이
- 평균 단어 길이(i.e., 문장 길이 / 단어 개수)

- 공백 문자와 단어를 제외한 쓰레기 문자의 개수
- spam과 ham에서 각각 많이 사용된 단어 내림차순 정렬

실험 결과, Fig. 2와 3번에서 볼 수 있듯이 spam은 ham과 달리 단어 개수 및 문장의 길이가 일정 임계치 주변의 값을 가지고 있다는 것을 알 수 있었다. 그림들에서 녹색 점은 spam을, 빨간색 점은 ham을 나타낸다. Fig. 2를 통해 spam문장의 단어 개수는 ham문장의 단어 개수보다 평균적으로 많고 특정 값 사이 (i.e., 25에서 35 사이)에 있다는 것을 알 수 있다. 또한, Fig. 3은 ham문장과 달리 spam문장은 일정 길이 이상을 가지고 있음을 보여준다. 일상적인 정상 문자 메시지는 상대방과 양방향 대화를 기반으로 서로 의미전달이 주요 목적이기 때문에 때로는 질문에 대한 간단한 응답 메시지도 생각보다 많이 발생해서 평균적인 문장의 길이가 짧음

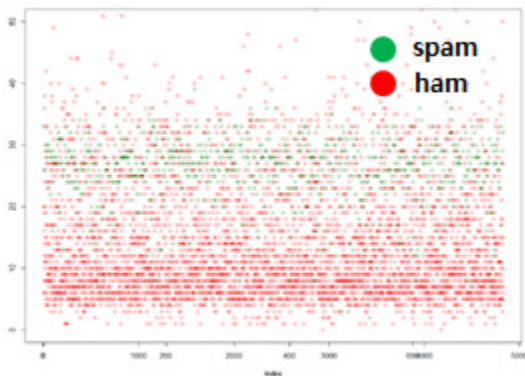


Fig. 2. Word counting comparison between spam and ham

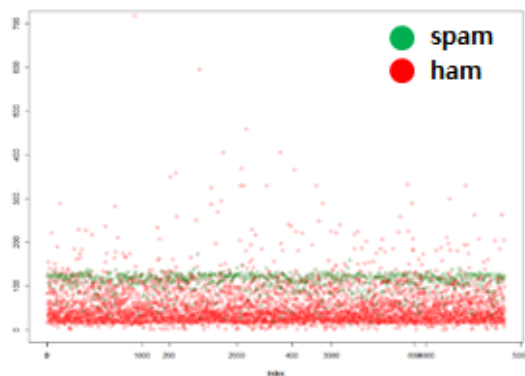


Fig. 3. Sentence length comparison between spam and ham

을 알 수 있다. 반면, spam은 일반적으로 문자 메시지를 보낼 때 가능하면 의도한 바를 충실히 전달해야 하기 때문에 평균적으로 ham보다 문장의 길이가 길다는 것을 알 수 있다.

추가적으로, Fig. 4는 spam과 ham에서 공백 문자와 단어를 제외한 특수 문자들의 개수 비교를 보여준다. 직관적으로 예상할 수 있듯이 spam문장에는 특수 문자들이 ham보다 많음을 알 수 있다.

즉, 기초 실험 결과 다음과 같은 결론을 내릴 수 있다.

- 스팸과 정상 문자 메시지 사이엔 확연한 차이가 보인다. 특히 스팸의 경우 스팸 문자를 구성하고 있는 단어 개수, 특수 문자 개수 및 스팸 문자 길이가 규칙성 있음을 알 수 있다.
- 이는, 스팸이란 일종의 정보를 제공해서 어떤 이득 또는 목적을 달성해야 하기 때문이다. 만약 아무런 정보를 제공하지 않는다면, 이는 굳이 스팸을 발송할 필요가 없기 때문이다.
- 반면, 일반적인 문장은 서로 주고받는 대화형이기 때문에 평균적으로 긴 문장일 필요가 없을 것으로 예상된다.
- 또한, 특정한 의도 및 목적 달성을 위한 spam도 있지만, 특수 및 쓰레기 문자들이 spam에 더 많은 것을 봤을 때 아무 의미 없는 단순한 쓰레기 문자들을 발송시키는 spam문자 메시지도 상대적으로 존재함을 알 수 있다.
- 기본 실험을 바탕으로 ham문장의 길이는 대략 100보다 훨씬 작은 범위에 몰려있음을 알 수 있다.

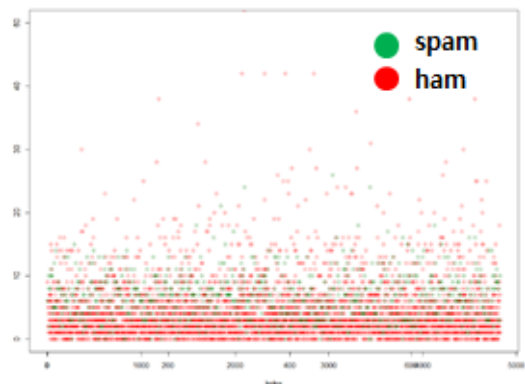


Fig. 4. Special letters counting comparison between spam and ham

3.2 확장 실험 결과

기초 실험 결과를 통해서 스팸문장과 정상문장이 기본적으로 차이가 있음을 보였다. 하지만, 이런 단순한 규칙으로 spam과 ham을 구분하는 데는 한계가 있다. 따라서 기초 실험에서 나온 결과를 바탕으로 스팸 의심 구역으로 접어드는 문장을 선택한 후, 단어의 다양한 특징들을 추가적으로 파악하여 적용해야 함을 알 수 있다. 예를 들어, spam으로 의심이 되는 문장에 대해 스팸에서 주로 등장한 단어의 빈도수 및 단어와 단어와의 관계 등을 추가적으로 고려해 볼 필요가 있다.

단어에 의미를 부여하기 위해서 spam과 ham에서 각각 많이 사용된 단어 26개를 Fig. 5와 같이 내림차순 정렬하고 비교해봤다. 문법 흐름상, be동사 및 일반 동사들 같은 경우 spam과 ham에 공통적으로 사용되는 경우가 많음을 알 수 있었다. 각 단어의 영향력을 살펴보기 위해, spam과 ham에서 사용된 공통단어들을 삭제하고 영향력이 상대적으로

is	129	is	588
call	98	m	284
FREE	86	get	281
f	78	have	277
claim	57	be	261
free	55	s	254
won	52	know	228
have	50	are	224
send	49	go	220
get	45	come	209
new	42	got	184
URGENT	41	do	176
contact	37	want	152
awarded	37	going	152
Get	33	good	148
Free	32	was	132
are	32	call	129
receive	30	need	118
mobile	30	think	115
apply	28	see	115
find	26	am	112
latest	26	don	107
guaranteed	26	send	107
selected	25	take	99
know	25	tell	92
WON	23	give	86
150p	23	ask	86

Fig. 5. Decreasing sorting comparison of all words between spam and ham

FREE	86	ask	86
f	78	doing	84
claim	57	said	80
URGENT	41	dear	63
awarded	37	pick	63
guaranteed	26	sure	61
WON	23	buy	60
150p	23	feel	56
GUARANTEED	23	went	55
weekly	20	told	53
Urgent	17	coming	51
national	17	Are	44
PRIVATE	16	fine	44
Valid	15	nice	43
top	12	few	43
charged	12	da	42
co	12	same	37
complimentary	12	sleep	36
redeemed	12	having	35
Claim	11	eat	35
valued	11	big	33
opt	10	work	32
Calls	10	dun	29
Latest	10	pay	28
Bonus	10	forgot	27
CALL	9	working	27
Offer	9	bad	26

Fig. 6. Decreasing sorting comparison of each independent word between spam and ham

높은 형용사와 동사만 남도록 한 단계 더 전처리를 수행했다. 결과, Fig. 6와 같이 spam 및 ham별 많이 사용된 단어들을 뽑아낼 수 있었다.

또한, Fig. 7은 각 단어가 ham에서 등장한 횟수와 spam에서 등장한 횟수를 도식화해서 그림으로 보여준다. 각 점은 각 단어를 의미한다. 예를 들어, is이라는 단어는 ham에서는 588번 등장했기 때문에 x축 588번 위치에 녹색 점이 존재하고, spam에서는 129번 등장했기 때문에 y축 129번 위치에 빨간색 점이 존재하게 된다. Fig. 7의 spam과 ham에서 사용된 단어들의 경향을 보면 한쪽으로 사향되어 (skewed property) 나타남을 알 수 있다. 이는 spam과 ham구분 시 영향력이 큰 단어들을 고려해야 함을 도식화 각도에서 다시 알려준다.

기초 및 확장 실험을 바탕으로 본 논문에서는 문장의 평균적인 길이, spam과 ham에서 각각 주로 사용되는 단어의 빈도수와 동시출현 단어분석 결과를 계층적으로 반영하여 spam 메시지와 ham 메시지를 구분해주는 기법을 제안한다.

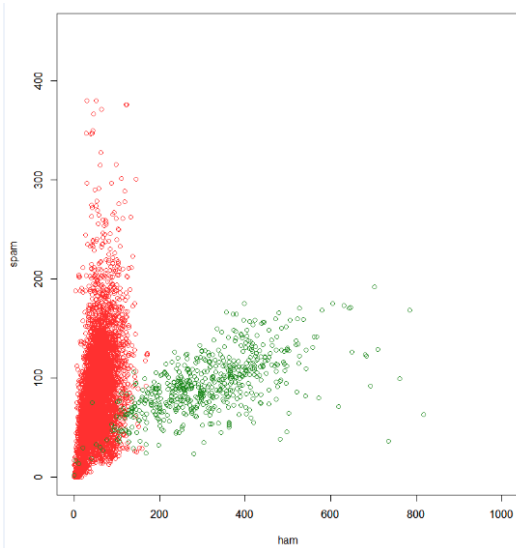


Fig. 7. Skewed properties of each word

IV. 동시출현 단어분석 기반 스팸 SMS 탐지 기법

본 논문에서는 동시출현 단어분석 기반 스팸 SMS탐지 기법을 제안한다. 이를 위해, 기초 및 확장 실험을 바탕으로 전처리 전 데이터를 구성하고 있는 spam과 ham을 구성하고 있는 각 단어들의 경향을 파악하고 단어와 단어 간의 연관성을 활용하여 동시출현 단어들을 분석한다. 즉, spam이든 ham이든 일반적으로 의미전달은 적어도 2개 이상의 단어가 구성될 때 발생하므로 본 논문에서는 한 문장을 구성하는 특정 단어와 단어들 간의 연관성을 파악하고 동시출현 단어분석 결과를 활용해서 spam 문자를 탐지한다.

4.1 단어와 단어의 동시출현 분석

두 개의 단어 조합이 각각 spam과 ham에서 얼마나 비중을 두고 있는지 고려하기 위해, Fig. 8과 같이 한 문장을 이루는 단어와 단어 사이의 가능한 모든 조합 규칙을 정의했다.

특정 단어와 단어의 조합이 spam에서 나타나면 spam에서의 빈도수를 1 증가시키고, ham에서 나타나면 ham에서의 빈도수를 1 증가시켰다.

식 (1)은 spam인지 ham인지 아직 판별되지 않은 하나의 새로운 문장을 구성하는 모든 단어들 (n)의 조합들 (예: $nC2 = n(n-1)/2$)을 각 타입에서

I am a student.	I am a boy.
I - am	I - am
I - a	I - a
I - student	I - boy
am - a	am - a
am - student	am - boy
a - student	a - boy

Fig. 8. Combination example of each word in the proposed scheme

각 경우별로 빈도수를 계산 후, 해당 문장을 구성하는 총 단어 수 (n)으로 나눈 결과를 보여준다.

$$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n f_{w_i, w_j}^T}{n} \quad (1)$$

where

w_i =each word in the sentence

w_j =each next word in the sentence

n= # of total words in the sentence

T = type classification (예: spam 혹은 ham)

예를 들어, Fig. 8의 문장 (예: I am a student)가 spam인지 ham인지 판단하려면 6가지 동시 단어 (예: I-am, I-a, I-student, am-a, am-student, a-student)의 각 조합들의 비율이 spam경향이 강한지 ham경향이 강한지 식 (1)을 바탕으로 계산하게 된다. 즉, 새로 고려하는 한 문장의 단어 대 단어 조합의 경향이 spam경향이 크면 spam으로 ham경향이 크면 ham으로 판별하게 된다. Fig. 9는 제안하는 기법의 탐지 기법의 정확도가 거의 98%이상 된다는 것을 보여준다. Fig. 9의 해석방법은 다음과 같다.

- 1번째 열은, 제안한 기법으로 판별된 결과를 나타낸다. 0 이면 잘못된 판단 (실제와 예측이 다른 경우), 1이면 정확한 판단을 한 경우 (실제와 예측이 같은 경우) 이다.
- 2번째 열은, 원본 데이터 셋에 적혀있던 타입이다.
- 3번째 열은, 제안된 기법을 통해 판단/예측한 타입이다. 원래 데이터의 2번째 열과 같은 결과가 나오면 제안하는 기법이 좋다는 것을 의미한다

다. 즉, 2와 3번째 열이 같으면 1번째 열이 1로 적힌다.

- 4번째 열은, 현재 고려하고 있는 새로운 문장의 총 단어 수 (n)이다.
- 5번째 열은, 원본 데이터 셋에서 ham으로 판명된 문장들에서 식(1)에서 고려하고 있는 단어 대 단어 쌍이 등장했던 평균값이다.
- 6번째 열은, 원본 데이터 셋에서 spam으로 판명된 문장들에서 식(1)에서 고려하고 있는 단어 대 단어 쌍이 등장했던 평균값이다.

```

1 success type map_type word_count ham_percent spam_percent
2 1 ham ham 20 2.370916 0.203005%
3 1 ham ham 6 0.297383 0.000000%
4 1 spam spam 33 2.501272 18.659159%
5 1 ham ham 11 0.929501 0.076287%
6 1 ham ham 14 1.323888 0.340630%
7 1 spam spam 36 9.434166 10.721006%
8 1 ham ham 16 1.991756 0.704745%
9 1 ham ham 26 26.066380 1.881958%
10 1 spam spam 26 2.144819 13.659048%
:
5566 1 ham ham 18 4.355257 0.117748%
5567 1 ham ham 5 0.148986 0.000000%
5568 1 ham ham 15 2.913060 1.759620%
5569 1 ham ham 3 0.039262 0.000000%
5570 1 spam spam 29 2.819098 7.042871%
5571 1 spam spam 32 2.962825 17.779692%
5572 1 ham ham 7 0.397065 0.140800%
5573 1 ham ham 10 1.539447 0.048443%
5574 1 ham ham 27 7.772684 2.481909%
5575 1 ham ham 6 0.405204 0.102436%
    
```

Fig. 9. Accuracy of the proposed scheme

V. 성능평가

제안하는 기법의 성능을 검증하기 위해서 NUS 데이터 셋을 5-fold validation해서 평가해왔다. 테스트 데이터 셋을 다양한 비중으로 활용하여 제안하는 기법을 성능평가 해본 결과 거의 98%이상의 정확도가 있음을 알 수 있었다.

Fig. 10은 NUS 데이터 셋에서 일반적인 spam

```

spam sports fans - get the latest sports news str+ 2 ur mobile
1 wk FREE PLUS a FREE TONE Txt SPORT ON to 8007
www.setzed.co.uk 0670141701215+ norm 4txt/120p
spam Urgent Urgent! We have 800 FREE flights to Europe to give away,
call 84 10th Sept & take a friend 4 FREE.
Call now to claim on 0905000555. BA128NNFWFLV150ppm
spam Great NEW Offer - DOUBLE Mins & DOUBLE Txt on best Orange tariffs
AND get latest camera phones 4 FREE! Call MobileUpd8 free on
08000839402 NOW! or 2stopxtxt T&Cs
spam Someone U know has asked our dating service 2 contact you!
Cant Guess who? CALL 09058097189 NOW all will be revealed.
POBox 6, L15HB 15Qp
spam Bored housewives! Chat n date now! 0671750.77.11!
BT-national rate 10p/min only from landlines!
spam We tried to call you re your reply to our sms for a video mobile
750 mins UNLIMITED TEXT free camcorder Reply or
call now 0800030705 Del Thurs
spam This is the 2nd time we have tried to contact u.
U have won the £ 400 prize. 2 claim is easy, just call
06710471148 NOW! Only 10p per minute. BT-national-rate
    
```

Fig. 10. Examples of spam sentences in the original data set

```

spam Hello darling how are you today?
I would love to have a chat,
why dont you tell me what you look like and what you are in to sexy?
0 spam ham 29 15.172976 10.325163
    
```

Fig. 11. Example of false negative sentence

문장들의 예시를 Fig. 11은 제안하는 기법에서 잘못 판정된 spam문장의 예를 보여준다. Fig. 11번의 spam문장의 단어들은 Fig. 10번의 일반적인 spam문장들과는 구조가 확연히 차이가 나고 오히려 ham문장에서 사용된 단어들로 구성되어있다. ham과 유사해보이는 몇몇의 spam문장들은 거짓음성 (false negative : FN)를 작은 비율로 야기함을 알 수 있다.

제안한 기법의 성능을 평가하기 위해 다양한 각도에서 거짓 양성 (false positive : FP)와 거짓 음성을 살펴보았다. 거짓 양성 통계상 실제로는 양성인데 검사 결과는 음성이라고 나오는 것이다. 예를 들어, 어떤 문자가 스팸 문자인지 검사한다고 할 때 어떤 문자가 실제로는 정상 문자지만 제안한 기법에서 검사 결과 스팸 문자라고 판정한다면, 이것이 거짓 양성이다. 반면, 거짓 음성은 실제로는 스팸 문자이지만 정상 문자라고 판단한 것이다. Fig. 12와 같이 training sentences개수를 변경하면서 실험해본 결과 제안하는 기법은 거짓 양성은 1.25%이하이고, 거짓 음성은 0.53%이하임을 알 수 있다. 또한, precision (TP/TP+FP)가 recall (TP/TP+FN)보다 낮은 값을 알 수 있다.

FP가 높으면 정상 문자를 스팸문자로 인식해서 정상 문자가 수신되지 않을 수 있기 때문에 사용자에게 불편함을 줄 수 있다. 반면, FN가 높으면 다양한 스팸 문자 서비스가 다수 전송되어 개인정보 유출 및 금전 피해 등까지 줄 수 있다. 따라서 가능하면 FN를 낮추는 것이 더 중요하다.

# of training sentences	% of training data set	# of errors
2787	50%	99
3344	60%	86
3901	70%	69
4459	80%	54
5016	90%	42
5574	100%	34

Fig. 12. False positive and false negative of the proposed scheme

VI. 결 론

본 논문에서는 동시출현 단어 조합을 활용하여 스팸 문자 탐지 기법을 제안했다. 공식적으로 공개된 스팸 문자 데이터 셋을 다 각도에서 전처리 실험해 본 결과, 스팸 문자는 단순히 한 단어 또는 간단한 문장으로 의미전달 하기는 어렵기 때문에 특정 문장 길이를 가지고 있음을 알 수 있었다. 또한, 특정 한 단어를 사용한다기 보다는 적어도 단어와 단어의 조합을 통해서 의도를 전달한다는 경향을 파악할 수 있었다. 제안하는 기법은 문장의 길이, 단어 사용 빈도 수 및 단어와 단어의 조합 등을 다 각도로 고려해서 스팸 문자를 탐지하기 때문에 탐지율이 거의 98%에 달함을 알 수 있었다.

References

- [1] T. M. Mahmoud and A. M. Mahfouz, "SMS Spam Filtering Technique Based on Artificial Immune System," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, pp. 589-597, Mar. 2012.
- [2] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, "Contributions to the study of SMS spam filtering: new collection and results," in *Proceedings of the 11th ACM symposium on Document engineering*, pp. 259-262, Sep. 2011.
- [3] J. M. Gómez Hidalgo, G. Cajigas Bringas, E. Puertas Sanz, and F. Carrero García, "Content Based SMS Spam Filtering," in *Proceedings of the 2006 ACM Symposium on Document Engineering, Amsterdam, The Netherlands*, pp. 107 - 114, Oct. 2006.
- [4] Gordon V. Cormack, José María Gómez Hidalgo and Enrique Puertas Sánz, "Feature Engineering for Mobile (SMS) Spam Filtering", *ACM SIGIR'07*, pp. 871-872, Jul. 2007.
- [5] Huang Jie, Huang Bei and Pu Wenjing, "A Bayesian Approach for Text Filter on 3G Network", In *Proceedings of the 6th International Conference on Wireless Communications Networking and Mobile Computing*, pp. 1-5, Sep. 2010.
- [6] Wuying Liu and Ting Wang, "Index-based Online Text Classification for SMS Spam Filtering", *Journal of Computers*, vol. 5, no. 6, pp. 844-851, Jan. 2010.
- [7] Xia Hu and Fu Yan, "Sampling of Mass SMS Filtering Algorithm Based on Frequent Time-domain Area", In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining*, pp. 548-551, Jan. 2010.
- [8] Zi Chu, Steven Gianvecchio, Haining Wang and Sushil Jajodia, "Who is tweeting on twitter: human, bot, or cyborg?", In *Gates et al.*, pp. 21-30, Dec. 2010.
- [9] Sarah Jane Delany, Mark Buckley and Derek Greene, "SMS Spam Filtering: Methods and Data", *Expert Systems with Applications*, vol. 39, no. 10, pp. 9899 - 9908, Dec. 2012.
- [10] Almeida, T.A., Gómez Hidalgo, J.M., Silva, T.P. "Towards SMS Spam Filtering: Results under a New Dataset", *International Journal of Information Security Science (IJISS)*, vol. 2, no. 1, pp. 1-18, Jan. 2013.

〈 저자 소개 〉



오 하 영 (Hayoung Oh) 정회원

2002년 2월: 덕성여자대학교 컴퓨터공학과 졸업

2006년 2월: 이화여자대학교 컴퓨터공학과 석사

2013년 2월: 서울대학교 컴퓨터공학과 박사

2010년 4월~2010년 10월: U.C. Berkeley 방문연구원

2013년 3월~2013년 8월: 서울시립대학교 연구교수

2013년 9월~현재: 숭실대학교 전자정보공학부 조교수

〈관심분야〉 소셜 정보망, 추천시스템, 무선 네트워크 및 비디오 스트리밍