

# 휴대폰 SMS를 위한 SVM 기반의 스팸 필터링 시스템

정회원 조인휘\*, 준회원 심혜택

## A SVM-based Spam Filtering System for Short Message Service (SMS)

Inwhee Joe\* *Regular Member*, Hye-Taek Shim *Associate Member*

### 요 약

휴대 전화는 이제 우리의 일상생활에서 없어서는 안 될 중요한 가전 기기로 자리 잡았다. 이러한 와중에 휴대폰에서 사용하는 문자 메시지 사용량 역시 꾸준히 증가하여 현재는 음성 통화 이용량의 1.5배에서 2배에 이르고 있다. 문자 메시지의 사용량이 증가함에 따라 스팸 문자 메시지도 따라서 증가 하였는데 기존의 모바일 기기에서의 스팸 필터링 방식은 단순 문자열 비교나 특정 번호 차단과 같은 아주 기초적인 수준으로 스팸 메시지를 필터링 하고 있는 실정이다. 본 논문에서는 SVM(Support Vector Machine)과 시소러스(thesaurus) 사전을 이용하여 좀 더 강력하고 적응적인 스팸 필터링 시스템을 제안하였다. 제안한 시스템은 샘플 문자 메시지로부터 전처리기를 이용하여 문자 메시지 속에 담겨 있는 단어를 추출 한 후, 추출된 단어를 시소러스 사전을 이용하여 해당 의미가 가지는 대표 단어로 변경하였다. 변경된 단어들에서 카이 제곱 통계량을 계산하여 그 값이 높은 단어들을 특징 단어로 선정하였고 선정된 특징 단어들을 가지고 SVM 분류기로 학습을 진행하였다. 그 후 학습된 분류기를 이용하여 테스트 문자 메시지의 스팸 여부를 분류 하였으며 평균 92%의 인식률을 보였다. 제안된 시스템은 PC에서 구현되어 있으며 실험을 통하여 그 성능을 확인하였다.

**Key Words** : Spam Filtering, SVM, Thesaurus, SMS, Cell Phones

### ABSTRACT

Mobile phones became important household appliance that cannot be without in our daily lives. And the short messaging service (SMS) in these mobile phones is 1.5 to 2 times more than the voice service. However, the spam filtering functions installed in mobile phones take a method to receive specific number patterns or words and recognize spam messages when those numbers or words are present. However, this method cannot properly filters various types of spam messages currently dispatched. This paper proposes a more powerful and more adaptive spam filtering system using SVM and thesaurus. The system went through a process of isolating words from sample data through pro-processing device and integrating meanings of isolated words using a thesaurus. Then it generated characteristics of integrated words through the chi-square statistics and studied the characteristics. The proposed system is realized in a Window environment and the performance is confirmed through experiments.

### 1. 서 론

휴대 전화는 이제 우리의 일상생활에서 없어서는 안 될 중요한 가전 기기로 자리 잡았다. 이러한 외

에 휴대폰에서 사용하는 문자 메시지 사용량 역시 꾸준히 증가하여 현재는 음성 통화 이용량의 1.5배에서 2배에 이르고 있다. 문자 메시지의 사용량이 증가함에 따라 스팸 문자 메시지도 따라서 증가 하였는데

\* 한양대학교 컴퓨터공학부 이동네트워크 연구실 (iwjoe@hanyang.ac.kr)

논문번호 : KICS2009-07-283, 접수일자 : 2009년 7월 6일, 최종논문접수일자 : 2009년 9월 14일

2004년 12월 1인당 평균 하루 1.7통에서 2005년 5월 0.6건으로 감소했으나 12월 조사에서는 0.74통, 2006년 3월에는 0.99통으로 다시 증가 추세를 보이고 있다. 이러한 이동 전화 스팸 메시지로 인해 개인이 느끼는 피해 유형을 보면 ‘불쾌감, 사생활 침해 등 정신적 피해’ (32.3%), ‘시간 낭비’ (24.8%), ‘개인 정보 유출’ (21.3%) 등 주로 정신적 피해를 많이 겪고 있는 것으로 나타나고 있다.

현재 휴대폰 내에서 이러한 스팸 메시지에 대해 필터링 방식은 특정 문자열이 문자 메시지에 있을 경우에 스팸으로 분류 한다거나, 특정 번호를 입력 하여 스팸 문자 여부를 판단하는 아주 기초적인 지원 중이다. 하지만 이러한 방식만으로는 현재 발송 되고 있는 스팸 문자 메시지에 대한 대처를 제대로 해주기 힘든 실정이다.

본 논문에서는 다양한 스팸 문자 메시지의 필터링을 위하여 SVM과 시소러스 사전을 이용하여서 내용 기반의 스팸 문자 메시지 필터링 시스템을 구축 하였고 최적의 성능을 내기 위한 다양한 방법을 연구 하였다.

논문의 구성은 다음과 같다. 1절은 본 연구의 배경과 필요성, 그리고 논문의 구성에 대해 설명하였다. 2절은 스팸 필터링 시스템의 설계 및 구현에 관하여 기술 하였으며, 3절은 시스템을 구축하여 실험한 과정을 정리하고, 4절에서 실험한 결과를 바탕으로 결과를 정리, 분석한 후 향후 연구 과제를 제시하였다.

## II. 설계 및 구현

### 2.1 전체 구성

제안하는 시스템의 전체 흐름은 다음과 같다.

1 단계: 다양한 휴대전화 사용자로부터 스팸/비 스팸 문자 메시지를 수집하고 수집된 문자 메시지를 학습 데이터와 테스트 데이터로 분류하는 단계.

2 단계: 수집된 데이터들을 스팸 여부와, 전화번호부에 등록 여부를 확인하여서 문자 메시지 포맷을 만드는 단계

3 단계: 학습 데이터로부터 특징 벡터를 추출 하는 단계

4 단계: 추출된 특징을 이용하여 SVM 최적 경계면을 학습 하는 단계

5 단계: 테스트 데이터로 실제 스팸 필터링 시스템을 테스트 하는 단계

시스템을 사용하기 전에 다양한 휴대전화 사용자로부터 스팸/비 스팸 문자 메시지를 수집하여서 이것을

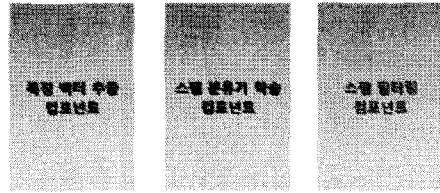


그림 1. 전체 시스템 구성도

학습 데이터와 테스트 데이터로 분류 한다. 문자 메시지 포맷은 다음과 같은 정보를 포함한다. 해당 문자가 스팸일 경우 1을 그렇지 않을 경우 0값을 저장하고 해당 문자 메시지의 수신 번호의 전화 번호부 등재 여부를 등재 되었을 경우 0 등재 되지 않았을 경우 1로 저장한다. 해당 문자 메시지의 본문과 위의 두 개의 정보를 합쳐 하나의 문자 메시지 데이터로 정의 한다.

제안하는 시스템은 그림 1에서 보듯이 크게 3개의 컴포넌트로 구성된다. 첫 번째 특징 벡터 추출 컴포넌트는 학습 데이터로부터 특징 벡터로 사용 할 단어들을 추출 하여 특징 벡터를 구성하는 컴포넌트이고 두 번째 스팸 분류기 학습 컴포넌트는 학습 데이터를 추출된 특징 벡터를 이용해서 SVM classifier의 최적 경계면을 수정 하는 컴포넌트이다. 스팸 필터링 컴포넌트는 실제로 스팸 문자의 여부를 판단하는 컴포넌트으로써 삼입된 문자 메시지가 스팸인지 여부를 판단한다. 스팸일 경우 1 스팸이 아닐 경우 0으로 판단 한다.

특징 벡터 추출 컴포넌트와 스팸 분류기 학습 컴포넌트는 학습 엔진으로써 PC환경에서 수행을 하도록 설계 하였으며 스팸 필터링 컴포넌트는 실제로 스팸 문자 메시지를 분류 하는 컴포넌트이므로 임베디드 환경에서도 동작 할 수 있도록 설계 및 구현을 하였다.

### 2.2 특징 벡터 추출 컴포넌트

특징 벡터 추출 컴포넌트는 학습 데이터로부터 특징 벡터를 추출하는 컴포넌트이다. 특징 벡터란 일종의 단어의 존재 유무를 0또는 1로 표현한 배열이라고 볼 수 있다. 학습 데이터에서 전 처리기를 통해서 단어를 추출하고 이것 들 중 변별력이 높은 단어를 특징으로 선정 한다. 다음 그림 2는 특징 벡터 추출 컴포넌트의 구조도이다.

특징 벡터 추출 컴포넌트는 학습 데이터를 입력 받으면 전처리 과정을 거쳐서 단어의 원형을 얻게 된다. 전처리 과정은 총 4단계로 구성되며 그 내용은 다음과 같다.

- 특수 문자 제거

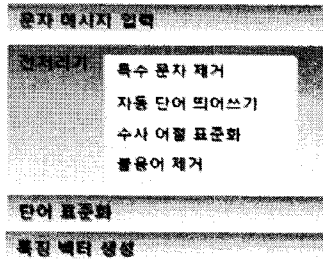


그림 2. 특징 벡터 추출 컴포넌트 구조도

- 자동 단어 띄어쓰기
- 수사 어절 표준화
- 불용어 제거

스팸 문자 메시지는 다량의 특수 문자를 포함 하는 경우가 많다. 이러한 특수 문자는 문자 메시지에서 학습하고자 하는 단어를 오인식 하게 하는 주된 원인이 된다. 그러므로 앞말은 특수 문자 제거를 함으로써 단어의 정상적인 추출을 유도 한다. 특수 문자가 제거된 문자 메시지는 단어의 띄어쓰기가 필요한 경우가 많다. 문자 메시지의 한계인 80byte 때문에 메시지가 때론 붙어서 오는 경우도 많고 특수 문자를 제거함으로 인해서 글자 들 사이에 빈 공간이 생겨서 단어를 제대로 인식 하지 못하는 경우도 많다. 이런 문제를 자동 단어 띄어쓰기 엔진을 이용해서 원래 의미의 단어로 복구하는 작업을 진행 한다.

수사 어절이란 1000원 혹은 천원 과 같은 숫자어절을 말하는 것이다. 수사 어절 표준화라는 것은 이러한 수사를 하나의 단일 표기로 표기 하는 것을 의미 한다. 수사 어절 표준화를 거치면 삽입된 문자 메시지에 실제로 수사가 있는지 여부를 알 수 있고 이 수사가 단순한 수의 표현인지 아니면 돈의 수치에 관련된 부분인지도 알아 낼 수 있다. 돈의 수치일 경우 특징 벡터에 그 정보를 표시 한다. 돈과 관련된 단어 여부까지 파악 하는 이유는 스팸 메시지의 상당량이 불법 대출에 관한 메시지이고 이런 불법 대출 메시지에는 대부분 대출액이 쓰여 있기 때문이다.

불용어란 전치사, 조사, 관사, 접속사 등과 같이 주제 색인어로서의 의미가 없는 단어들을 말한다. 이러한 불용어를 제거해서 단어의 원형을 취함으로써 단어의 빈도수를 올바르게 확인 할 수 있다. 단어의 빈도수는 특징 벡터를 추출 할 때 카이제곱 통계량에서 필요하다. 불용어 처리를 통한 단어의 원형은 전처리기의 핵심이라고 볼 수 있다.

전처리기를 통과한 학습 데이터는 시소러스 사전을

이용한 단어 표준화 작업을 거친다. 단어 표준화란 비슷한 의미를 가지는 단어를 하나의 단어로 간주 하는 작업을 의미 한다.

각각의 학습 데이터들은 전처리기를 거쳐 표준화된 단어들의 집합으로 구성이 되는데 이 때 각각의 단어 들에 대해서 의미가 비슷한 단어에 대해 동일한 단어로 변환을 시켜주는 작업을 진행 한다.

예를 들면 TV나 텔레비전은 같은 의미를 가지지만 다른 표현을 가지게 된다. 만약에 이런 단어들이 있을 경우 각각의 단어들을 합쳐주지 않을 경우에는 TV라는 단어가 1의 빈도수를 가지고 텔레비전이라는 단어가 1의 빈도수를 가지게 되지만 이 두 개의 단어를 합쳐서 텔레비전이라는 단어로 치환 하게 되면 빈도수는 2가 된다. 이러한 작업을 시소러스를 이용하여 진행 하도록 한다.

단어 표준화를 거친 학습 데이터들은 카이제곱 통계량을 계산 하여서 일정 개수의 특징 벡터를 선정 한다. 즉 문자 메시지들을 전처리기 및 단어 표준화를 거쳐서 모아 놓으면 학습 데이터 전체를 이용해서 카이제곱 통계량을 계산 한다. 카이제곱 통계량이 높을 수록 해당 단어의 특징으로서의 의미가 큰 것이다. 본문에서는 100, 150, 200, 300개의 특징 벡터를 설정 하였고 각각의 특징 데이터 개 수 별로 학습을 진행 하였다. 카이제곱 통계량이 높은 단어들을 특징으로 선택 하여서 스팸 분류기 학습 컴포넌트에서 사용한다.

### 2.3 스팸 분류기 학습 컴포넌트

스팸 분류기 학습 컴포넌트는 이전의 특징 벡터 추출 컴포넌트를 이용하여 추출된 특징을 이용해서 SVM 분류기의 최적 경계면을 만드는 컴포넌트이다. 스팸 분류기 학습 컴포넌트의 구조는 다음 그림 3과 같다.

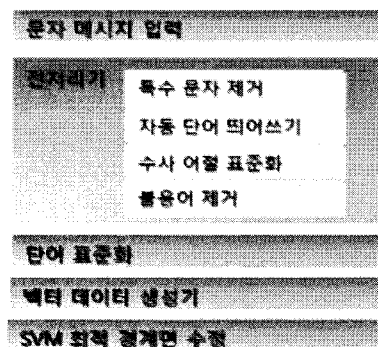


그림 3. 스팸 분류기 학습 컴포넌트

학습 문자 메시지가 입력되면 스팸 분류기 학습 컴포넌트는 이전의 특징 벡터 추출 컴포넌트와 동일한 방식으로 전처리기와 단어 표준화를 거친다. 이 두 단계를 거치고 난 문자 메시지로 벡터 데이터를 만들게 되는데 현재 삽입된 문자 메시지의 단어 들 중 특징 벡터와 일치 하는 단어가 있는지 확인을 하고 있을 경우 해당 값에 1을 표기 한다. 그리고 입력 된 문자 메시지의 연락처가 주소록에 있는지 여부를 확인 하여서 있을 경우 1을 없을 경우 0을 설정 한다. 마지막으로 수사 어절 표준화를 통해서 얻은 정보를 이용해서 문자 메시지 내용 중에 돈이 있을 경우 1을 없을 경우 0을 표기 한다.

그러한 형식으로 모든 특징 벡터를 0 또는 1로 값을 저장 한 후 SVM 분류기를 통해 학습을 진행 하도록 한다. 커널 함수로는 가우시안 RBF(Radial Basis Function) 커널 함수를 이용 하였다. 본 논문에서는 Constant 값을 10, 20, 40 의 세 단계로 학습 하였고 Gamma 값은 0.01, 0.05, 0.1의 세 단계로 하였다.

### 2.4 스팸 필터링 컴포넌트

스팸 필터링 컴포넌트의 구조는 다음 그림 4와 같다.

스팸 필터링 컴포넌트는 실제로 삽입된 데이터가 스팸인지 아닌지 판별 하는 컴포넌트이다. 이전 스팸 분류기 학습 컴포넌트에서 생성된 SVM 분류기를 이용하여서 스팸 여부를 분류 하는 것이 이 컴포넌트의 주된 목적이다.

입력된 문자 메시지는 전처리기와 단어 표준화를 통해서 단어를 추출 하였다. 하나 시소러스 사전의 경우 MySQL DB를 이용하여 구축 하였는데 휴대폰 환경에서는 MySQL과 같은 DBMS를 사용 할 수 없다. 고로 특징 벡터의 유의어만을 분리해 내어서 작은 규모의 시소러스 사전 파일을 구성하고 그것을 이용하

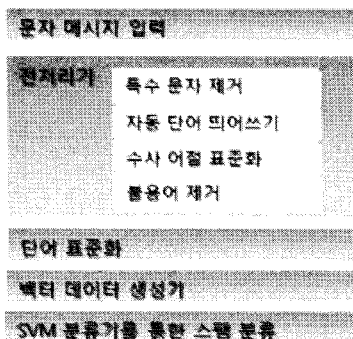


그림 4. 스팸 필터링 컴포넌트 구조도

여 단어 표준화를 진행 하였다. 추출된 단어들과 수사 정보, 전화 번호 정보를 통해서 벡터 데이터를 생성한 후 SVM 분류기에 해당 벡터 데이터를 넣으면 스팸 여부를 알 수 있다. 0 일 경우에는 비 스팸 문자 메시지 이고, 1일 경우에는 스팸 문자 메시지로 분류 된 것이다.

## III. 실험 및 결과 분석

### 3.1 스팸 필터링 컴포넌트

본 논문에서 제안한 시스템에 대한 성능 평가를 위해서는 문서 분류 및 정보 검색 분야에서 일반적으로 사용되고 있는 분할표를 이용하는 방법을 사용한다. n 개의 이진 결정을 내리는 시스템에서 n개의 결정 결과는 표 1과 같이 요약된다. 여기서  $n = a + b + c + d$ 가 된다<sup>7)</sup>.

표 1. 분할표

	시스템이 분류한 스팸 메시지	시스템이 분류한 일반 메시지
실제 스팸 문자 메시지	a	c
실제 일반 문자 메시지	b	d

$$SP = \frac{\text{시스템에 의해 맞게 분류된 실제 스팸 메시지의 수}}{\text{시스템에 의해 분류된 스팸 메시지의 총수}}$$

$$= \frac{a}{a+b} \quad (\text{if } a + b > 0)$$

$$SR = \frac{\text{시스템에 의해 맞게 분류된 실제 스팸 문자 메시지의 수}}{\text{실제 메시지의 총수}}$$

$$= \frac{a}{a+c} \quad (\text{if } a + c > 0)$$

여기서 SP는 스팸 정확률(Spam Precision)로 시스템이 스팸이라고 분류 한 메시지 중에서 올바르게 스팸을 분류한 메시지의 비율이고 SR은 스팸 재현율(Spam Recall)로 실제 스팸 메시지에서 올바르게 스팸 메시지를 예측 한 수치이다.

$$NSP = \frac{\text{시스템에 의해 맞게 분류된 실제 일반 문자 메시지의 수}}{\text{시스템에 의해 분류된 스팸 문자 메시지의 총수}}$$

$$= \frac{d}{c+d} \quad (\text{if } c + d > 0)$$

$$NSR = \frac{\text{시스템에 의해 맞게 분류된 실제 일반 문자 메시지의 수}}{\text{실제 일반 문자 메시지의 총수}}$$

$$= \frac{d}{b+d} \quad (\text{if } b + d > 0)$$

NSP는 비 스팸 정확률(Non-Spam Precision)로 시스템에 의해 스팸 문자 메시지가 아니라고 분류된 문자 메시지 중 실제로 스팸 문자 메시지가 아닌 비율이며 NSR은 비 스팸 재현율(Non-Spam Recall)로 실제 비 스팸 문자 메시지의 총 개수 중 비 스팸 문자 메시지로 올바르게 분류된 비율을 의미한다.

분할표의 수치 중 비 스팸 문자 메시지로 분류되는 비율인 b의 값과 스팸 메일이 일반메일로 분류되는 c의 값을 줄이는 것이 가장 중요하다. 성능 측정 중 모든 메일을 스팸으로 간주해 버리는 단순한 분류기의 경우 SR(스팸 재현률)은 100%를 나타내지만 SP(스팸 정확률) 매우 낮게 된다. 그와는 반대로, 모든 메일을 일반 메일이라고 간주하는 필터링 시스템은, 100%의 NSR(비스팸 재현률)을 보이지만 NSP(비스팸 정확률)는 매우 낮게 된다. 일반적으로 스팸 필터링 시스템은 내부 설정에 따라 재현률과 정확률 사이에 이율배반을 보인다.

스팸 문자 메시지 필터링 시스템의 성능 평가에 대한 실험은 다음과 같은 환경에서 이루어졌다. 시스템은 Windows 운영 체제에서 구축되었다. 스팸 문자 메시지 필터링 엔진은 C언어를 이용하여 만들어졌으며 시소러스는 MySQL DB를 이용하여 구축하였다. 전처리기는 한국어 형태소 분석기를 [6] 이용하였다. 이 한국어 형태소 분석기를 통해 형태소 분석 및, 불용어 제거, 자동 단어 띄어쓰기, 수사어절 표준화를 진행하였다. SVM 학습의 도구로는 OpenCV의 SVM모듈을 이용하였다<sup>7)</sup>. OpenCV는 오픈 소스 기반의 영상 처리를 위한 컴퓨터 라이브러리로써 KNN, SVM, 결정 트리, 신경망 등 다양한 기계 학습 엔진을 제공한다.

### 3.2 실험 결과

학습에 사용한 문자 메시지는 스팸 문자 메시지 100개, 비 스팸 문자 메시지 200개이며 실제 테스트에 사용된 스팸 문자 메시지는 80개, 비 스팸 문자 메시지는 80개이다. 실험은 크게 두 가지 종류로 나뉘는데 첫 번째는 특징 벡터 수 별 인식률의 결과이며 두 번째는 Gamma값과 Constant값에 따른 인식률 결

표 2. 특징 벡터 수 별 인식률 결과

벡터수	SP	SR	NSP	NSR
100	95.89%	87.5%	88.5%	96.25%
150	93.58%	91.25%	91.46%	93.76%
200	93.24%	86.25%	87.2%	93.76%
300	86.66%	81.25%	82.35%	87.5%

과이다.

카이제곱 통계량으로 단어들의 상대 도수를 분석한 후 그 값이 높은 순으로 특징 벡터를 선정한 다음 테스트 데이터를 통해서 인식률을 비교한 것이다. 위의 표 2에서 알 수 있듯이 특징 벡터의 개수가 150개일 때 가장 안정적인 인식률을 보인다. 고로 Constant와 Gamma값에 따른 인식률에 대한 결과는 특징 벡터의 개수를 150개로 선정한 상태에서 실험을 진행하였다.

표 3에서 보듯이 특징 벡터 150개를 기준으로 한 Constant와 Gamma값에 따른 결과를 보면 최적의 Constant값은 20이고 그에 따른 Gamma값은 0.01일 때 가장 높은 인식률을 보였다.

표 3. Gamma와 Constant 값에 따른 인식률 결과

Constant (Gamma)	SP	SR	NSP	NSR
10 (0.01)	85.71%	75.02%	77.77%	87.5%
10 (0.05)	89.33%	83.75%	84.70%	90.15%
10 (0.1)	87.83%	81.25%	82.55%	88.75%
20 (0.01)	91.56%	94.98%	94.80%	91.25%
20 (0.05)	88.46%	86.25%	86.58%	88.75%
20 (0.1)	91.78%	83.75%	85.05%	92.50%
40 (0.01)	85.13%	78.75%	80.23%	86.25%
40 (0.05)	87.67%	80.02%	81.60%	88.75%
40 (0.1)	84.93%	77.50%	79.31%	86.25%

## IV. 결론

본 논문에서 제안한 스팸 문자 메시지 필터링 시스템은 문자 메시지가 도착하면 수신된 문자 메시지의 발신자와 그 내용을 이용하여 자동으로 스팸 문자 메시지를 분류하는 기능을 가지고 있다. 제안된 시스템은 특징 벡터의 개수가 150개일 때 평균 92%라는 만족할 만한 인식률 결과를 얻을 수 있었다.

제안한 스팸 필터링 시스템은 학습 데이터와 유사한 스팸 문자 메시지의 경우 높은 빈도수로 문자 메시지를 인식하나 학습되지 않은 패턴의 스팸 문자 메시지인 경우에는 검출률이 떨어지는 한계를 보였다. 이것은 기계 학습 알고리즘의 한계로서 다양한 패턴의 학습 데이터를 보유함으로써 극복이 가능할 것으로 예상된다. 그리고 동음이의어가 시소러스 단어 사전을 이용해서 유의어를 검색할 경우 다른 의미로 검색되는데, 이로 인해 오인식이 발생하였다. 이러한 동음이의어에 대한 처리는 단어의 전후 문맥을 분석하

여 해당 단어의 의미를 추정 하는 연구를 진행함으로써 동음이의어에 대한 오인식률을 낮출 수 있을 것으로 생각 된다.

참 고 문 헌

- [1] 임혜영, "SVM 분류기를 이용한 문서 범주화 연구", 연세대학교 문헌정보학과 석사학위 논문, 2000.
- [2] 박진우, 고영중, 서정연, "문서 요약 기법을 이용한 자동 문서 범주화", 제 13회 한글 및 한국어 정보처리 학술대회, 138-145, 2001.
- [3] C. Cortes and V. Vapnik, "Support vector network" Machine Learning, vol. 20, pp.273-297, 1995.
- [4] N. Cristianini, J. S. Talor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press 2000.
- [5] Berges, C. J. "A tutorial on Support Vector Machine for pattern recognition", Data Mining and Knowledge Discovery 2, pp121-167 1998.
- [6] 김태희, "스팸 메일 필터링 시스템에서 어휘 정보와 시소러스의 영향 분석", 대구대학교 컴퓨터 공학과 박사 학위 논문 2005.
- [7] 한국어 형태소 분석기 <http://nlp.kookmin.ac.kr/HAM/kor/index.html>
- [8] Yihui Xie "An Introduction to Support Vector Machine and Implementation in R". May 8, 2007.

조 인 휘 (Inwhee Joe)

정회원



1983년 2월 한양대학교 전자공학과 졸업

1994년 12월 미국 University of Arizona, Electrical and Computer Engineering, M.S.

1998년 9월 미국 Georgia Tech, Electrical and Computer Engineering, Ph.D.

1992년 12월 (주) 데이콤 종합연구소 선임연구원  
 2000년 6월 미국 Oak Ridge 국립연구소 연구원  
 2002년 8월 미국 Bellcore Lab (Telcordia) 연구원  
 2002년 9월~현재 한양대학교 컴퓨터공학부 부교수  
 <관심분야> Mobile Internet, Cellular System and PCS, Sensor Networks, Mobility Management