

# 사용자 이메일 반응 분포 계산과 사용자 그룹 스팸 메일 필터 필요성 점검

김종완\*, 남인길\*

\*대구대학교 컴퓨터·IT공학부

e-mail: jwkim@daegu.ac.kr

## Checking Spam Mail Filter Need for Each User Group from Computing Distribution of User Email Responses

Jongwan Kim\*, In-Gil Kim\*

\*School of Computer and Information Technology, Daegu University

### 요 약

본 연구에서는 이메일 사용자별로 제공받은 사용자 선호 정보 대상으로 EM 클러스터링을 수행하여 사용자 클러스터를 만든 후, 사용자 클러스터들의 이메일 반응 분포를 계산함으로써 사용자 취향에 따라 동일한 이메일에 대해서도 서로 다른 반응을 가질 수 있다는 사실을 확인하려고 한다. 그 결과로부터 현재의 내용기반 방식과는 다르게 사용자 선호도를 고려한 스팸 메일 필터 구축 방법을 제안한다.

### 1. 서론

수년전에 스팸 메일은 단순히 귀찮은 일이었으나, 오늘날에는 사용자 생산성과 시스템 안정성에 심각한 위협을 주는 스파이 또는 피싱(phishing) 기능을 수행하고 있다[1]. 이런 스팸 필터링 문제는 문서 분류 문제의 특별한 종류로서 볼 수 있다. 다양한 기계학습 알고리즘들인 나이브 베이저안 분류기(NBC: naive Bayesian classifier)와 서포트 벡터 머신(SVM: support vector machine)은 메타데이터 이메일 분류 작업을 위해 사용된 대표적인 방법들이다[2,3]. 이전의 스팸 메일 연구로부터, 우리는 NBC와 SVM 사이에는 내용기반 스팸 필터링에서 큰 차이가 있음을 알 수 있다.

메일 서버가 모든 사용자 대상으로 동일하게 스팸 여부를 판정하기 보다는, 실생활에서 이메일에 대한 사용자 반응은 개인 취향이나 선호에 따라 다를 수 있다. 그러므로 개인 취향에 기반한 사용자 중심 스팸 필터링 서비스를 제공하는 것은 의미있는 일이다. 본 논문에서는 대학생 그룹으로부터 사용자 취향과 이메일 반응을 수집하고 연관성 분류 마이닝 프로세스를 거쳐서, 정형 언어로 사용자 선호 온톨로지(user preference ontology)를 정의하는 규칙들을 생성한다. 이러한 온톨로지가 스팸 대응 시스템에 사용되는 과정을 보이기 위하여, 사용자 선호 기반 스팸 메일 필터 구축 방법을 설명한다.

### 2. 데이터 수집 및 전처리

본 장에서는 사용자 프로파일 설계 및 데이터 수집을 포함한 데이터 전처리 과정을 기술한다. 특히 사용자들은

샘플 이메일에 대한 반응을 결정하기 위하여 사용자 선호 정보뿐만 아니라 이메일 카테고리 정보를 함께 고려하므로, 사용자 선호 온톨로지 구축에 필요한 이메일 카테고리 정보를 추출하는 방법도 설명한다.

이메일들에 대한 사용자들의 반응을 조사하기 위해서 사용자 선호와 이메일에 대한 사용자 반응 부류를 표현하도록 사용자 프로파일(user profile)을 설계한다. 상용 웹메일 소프트웨어들은 사용자의 개인 정보를 요구한다. 이때 요구되는 사용자의 등록 정보처럼, 본 논문은 사용자 프로파일에 포함할 특성들로 15가지 {Age, Gender, RequiredHits, Adults, Entertainment, Finance, Jobs, Shopping, Travel, News, Sports, Kids, IT, RealEstates, Etc}를 선택하였다. 나이(Age) 속성은 실험 참여자가 대학생이므로 두 그룹 FS(1-2학년)과 JS(3-4학년)으로 나누었다. RequiredHits (지금부터 RHit로 표기)는 어떤 메일이 스팸으로 간주되기 전에 스팸과 관련된 용어들이 본문에 얼마나 많이 나타나는지를 의미하는데, 이메일의 내용을 고려하지 않고 언어 표현이 사용자에게 보다 편하기 때문에 약함(W), 보통(N), 강함(S)의 언어 향으로 구별하였다. 만약 사용자들이 강한 스팸 필터를 원한다면, RHit 값으로 약함(W)을 선택할 것이고, 약한 필터를 원할 때는 강한(S) RHit를 선택하면 된다. 우리는 사용자 선호도가 얼마나 반응에 영향을 미치는지를 연구하기 위해, 사용자 선호도, 이메일 반응과 함께 기계학습에 의해 분류된 이메일 카테고리(이하 ECat) 정보도 데이터마이닝 프로세스에서 이용한다.

이메일 수신자들은 받은 이메일에 대해 보통 4가지, 즉

Reply(응답), Delete(삭제), Store(보관), Spam(스팸)으로 반응한다. 일부 사용자들은 스팸 메일을 바로 삭제하고 스팸 편지함으로 이동시키지만, 또 다른 부류의 사용자들은 스팸 메일들을 삭제만 하거나 이동만 시키는 단일한 행동을 수행한다. 분명한 사실은 대부분의 사용자들은 자신의 편지함에 스팸이 아닌 메일만 보관하기를 원한다는 점이다.

위의 사실에 근거하여 샘플 메일들, 실험에 참여한 사용자들의 개인 정보와 취향, 샘플 메일에 대한 반응들을 수집하였다. 개인 정보에 해당하는 속성들인 Age는 {FS, JS}로, Gender는 {male(M), female(F)}인 이진 형태로 주어지며, 대부분의 다른 취향 속성들도 이진 형식으로 되어 있다. 예를 들어, 만약 사용자가 Sports에 관심이 있다면 그 속성에 대해 true(T)를 선택하며 관심을 가지고 있지 않으면 false(F)를 선택한다. 반면에, RHit에서는 다진 속성값 (W, N, S)을 가지고, ECat 속성은 다른 이메일 카테고리 나타내기 위해 Adults부터 RealEstates까지의 11가지 카테고리 속성과 카테고리가 불분명한 기타(Etc) 속성을 포함해서 12가지 값을 사용한다. 사용자 반응 분류의 목표 변수 또는 출력 변수인 Response는 4가지 범주형 값인 (Reply, Delete, Store, Spam)을 가진다.

주어진 이메일에 대한 사용자 반응은 사용자 취향과 이메일 내용에 기반한 카테고리에 의존하므로, 이메일 카테고리 정보가 스팸 메일 필터에 포함되어야 한다. 일반적으로 성인, 금융, 취업 같은 카테고리로 이메일을 분류하는데 NBC나 SVM이 사용된다. 1장에서 언급했듯이, SVM은 이런 종류의 분류에 유용한 접근법이므로 본 연구에서는 이메일 카테고리 분류기로서 Vapnik이 제안한 표준 SVM[3]을 사용한다.

### 3. 사용자 그룹별 이메일 반응 분포 계산

본 연구의 수행 동기이자 기본 가정인 「사용자 선호 정보가 이메일 카테고리 정보와 함께 입력 이메일들에 대한 개별 사용자의 반응에 영향을 미칠 수 있다」의 유효성을 확인하기 위하여, Witten와 Frank[4]가 보급한 잘 알려진 데이터마이닝 개발 도구인 WEKA의 EM 클러스터링을 수집 데이터에 대하여 수행하였다. 기대-최대로 불리는 EM(expectation maximization) 알고리즘은 1977년에 A. Dempster 등이 제안하였으며, 기계학습과 컴퓨터비전, 자연어처리 등에서 광범위하게 사용되는 클러스터링 알고리즘으로서, 관찰되지 않고 잠재된 변수들에 의존하는 확률 모델에서 가장 큰 기대치를 갖는 파라미터 추정값을 발견할 목적으로 통계학에서 사용된다[5]. 실험에 참여한 74 사용자의 11가지 속성이 포함된 선호 레코드들 대상으로, EM 클러스터링을 수행한 결과 단지 2개의 클러스터가 최적으로 발견되었다.

본 연구에서는 사용자 선호 정보를 대상으로 클러스터화 된 사용자 그룹의 이메일 반응 분포를 계산하여 본 연구 수행 동기를 점검한다. 표 1은 각 이메일 카테고리별로

개별 반응인 Spam, Delete, Store, Reply의 수를 전체 반응수로 나눈 개별 분포를 보여준다. 대부분의 이메일들은 성인(adults), 오락(entertainment), 금융(finance), 취업(jobs), 쇼핑(shopping), 여행(travel) 및 기타(etc)의 7가지 카테고리로부터 얻어졌다. 어떤 개별 클러스터안의 각 사용자 선호에 대한 사용자 반응의 4가지 백분율 합이 가끔 100%가 되지 않는 이유는 소수의 이메일에 대하여 일부 사용자들이 자신의 반응을 제공하지 않았기 때문이다. 표 1에서 보듯이, 클러스터1과 클러스터2의 Spam, Delete, Store, Reply 반응은 이메일에 대한 사용자 선호도 정보 기준으로 서로 다르므로, 계산 결과는 개별 사용자 그룹이 사용자 선호에 따라 사용자 반응에서 다른 분포를 갖는다는 것을 확인시켜 준다. 최종적으로 이 연구의 동기인 사용자의 주관적인 선호도가 주어진 이메일과 이메일 카테고리 정보를 고려할 경우 사용자 반응에 영향을 미친다는 사실이 타당함을 입증하였다.

표 1. 주관적 사용자 선호도에 따른 사용자 클러스터별 사용자 반응 분포 (Spam(Sp), Delete(De), Store(St), Reply(Re))

		클러스터 #1 사용자 반응				클러스터 #2 사용자 반응			
		Sp	De	St	Re	Sp	De	St	Re
주관 적인 사용 자 선호 도(%)	성인	67.3	18.7	10.8	3.2	85.7	14.0	0.4	0
	오락	12.3	54.7	29.5	3.54	11.3	56.7	30.7	1.3
	금융	54.0	22.7	15.5	7.8	33.7	51.2	12.7	2.4
	취업	16.8	33.0	29.6	20.7	3.9	40.2	52.0	3.9
	쇼핑	19.1	48.6	29.4	2.9	14.0	55.7	28.4	1.7
	여행	12.2	43.1	40.4	4.4	17.3	41.3	38.3	3.1
	기타	25.8	45.0	18.3	10.9	11.7	61.7	15.6	7.5

### 4. 사용자 선호도 고려 스팸 메일 필터 구축

사용자 선호도를 고려한 스팸 메일 필터를 구축하기 위해서, 사용자 선호 온톨로지를 만든다. 온톨로지 생성은 본 연구자의 이전 연구[6]에서 설명된 바와 같이 3단계로 수행한다. 첫 번째 단계는 사용자들의 선호 정보, 이메일 카테고리 정보, 사용자들의 이메일 반응 정보 사이의 연관성 규칙들을 찾기 위해 연관성 분류 마이닝을 수행한다. 우리는 대표적인 의사결정 트리(decision tree) 알고리즘인 ID3[7]을 선택하였다. 샘플 데이터들 대부분은 이진 속성을 가지고 일부 범주형 속성들로 구성되어 있으므로, ID3가 데이터 집합으로부터 대표 규칙들을 발견하기에 적합하다. ID3 마이닝을 수행하여 의사결정 트리를 생성시킨 후, 트리의 각 경로를 하나의 규칙으로 기술함으로써 의사결정 트리를 규칙들로 변환한다.

다음 단계에서는, 불필요한 규칙들을 제거하고 이해하기 쉬운 규칙들만 남기기 위해 규칙 최소화를 수행한다

[6]. 카르노 맵(Karnaugh Map)은 불리언 논리 단순화를 이해하기 위한 간단하고 쉬운 방법으로 잘 알려져 있다. 카르노 맵에서는 동일한 함수에 대하여 2개 이상의 단순화된 논리 표현들을 발견하는 것이 가능하다. 적용된 규칙 최소화 방법은 불리언 변수와 범주형 변수들을 함께 다룰 수 있기 때문에 혼합형 논리합성 기반 규칙 가지치기 (hybrid logic synthesis based rule pruning: HLSRP)라 부른다. 데이터 마이닝으로부터 유도된 규칙 집합에는 몇 가지 변수들이 있다. 대부분의 변수들은 불리언 또는 이진 변수이지만, RHit와 같은 다진값을 가지는 변수도 있다. 혼합 논리합성에서는, 만약 어떤 변수에 대한 두 개의 필터링되는 논리가 다르다면, 그 변수를 가진 두 개의 규칙들은 하나로 병합되고 해당 변수에 있던 전제조건은 병합 규칙에서 생략되므로 간단한 규칙들이 유도될 수 있다. 아래의 예는 제안된 아이디어를 설명해준다.

R1: if Age=JS and RHit=Strong and News=True and Adults=False and Entertainment=False then Response>Delete.

R2: if Age=JS and RHit=Strong and News=True and Adults=False and Entertainment=True then Response>Delete.

두 규칙에서 Entertainment 속성의 전제조건만 다르므로, 논리합성 연산에 의해 Entertainment=True는 Entertainment=False와 결합하기 때문에 변수 Entertainment의 전제조건인 False와 True가 Null로 병합되고, 결국 두 개의 규칙 R1과 R2는 Null 조건을 배제한 새로운 규칙 R3으로 대체된다.

R3: if Age=JS and RHit=Strong and News=True and Adults=False then Response>Delete.

마지막 단계에서 최소화 규칙들을 도메인 온톨로지 안에 있는 규칙들로 변환하는 작업이 이루어진다. 온톨로지 (ontology)란 “실세계 혹은 특정 도메인에 존재하는 모든 개념들과 그 개념들의 속성들, 그리고 개념들이 상호간에 의미적으로 어떻게 관련되는지에 대한 정보를 가지고 있는 지식베이스”라 정의할 수 있고, 지능 시스템에서 정보의 의미를 정의하는데 중요한 역할을 수행한다[8]. 사용자 선호 기반 스팸 메일 필터에서는 사용자의 취향을 기반으로 한 사용자의 행동양식을 형식적으로 정의한 도메인 온톨로지가 도움이 될 수 있다. 분류 규칙들을 온톨로지와 온톨로지 간의 매핑을 표현하는데 적합한 강한 형식의 일차 논리 언어(strongly-typed first-order logic language)인 Web-PDDL[9]을 사용하여 온톨로지로 표현한다. 우리는 이전의 결과들을 기초로 Web-PDDL로 사용자 선호 온톨로지를 정의할 수 있다.

## 5. 결론

본 연구에서는 이메일 사용자별로 제공받은 사용자 선호 정보 대상으로 클러스터링을 수행하고 몇몇 사용자 그

룹의 이메일 반응 분포를 계산함으로써 사용자 취향에 따라 동일한 이메일에 대해서도 서로 다른 반응을 가질 수 있다는 사실을 확인하였다. 현재의 내용기반 스팸 메일 필터는 동일한 이메일에 대해서 같은 반응을 보일 것이라고 가정하는 범용 필터링 방식을 채택하고 있다. 하지만 우리는 사용자별 개인 성향과 이메일 반응 행태를 고려함으로써 사용자 선호도를 반영한 스팸 메일 필터 구축 방법을 제안하고 있다. 스팸 메일 필터 구축 과정에서 본 연구자가 선행 연구[6]에서 제안한 데이터마이닝 기술과 논리합성 방법을 이용하여 사용자 선호 온톨로지를 구성하는 방법을 활용하였다. 향후에는 현재의 내용기반 필터들과 비교하기 좋도록 사용자들에게 제공되는 이메일을 실시간으로 처리하도록 확장하고, 사용자 위주의 보다 다양한 스팸 메일 평가척도를 개발할 필요가 있다.

**감사의 글:** 이 논문은 2008년도 대구대학교 PoP-iT 누리사업단의 연구비를 일부 지원 받았음.

## 참고문헌

- [1] P. Wolfe, C. Scott, and M. Erwin, Anti-Spam Tool Kit, McGraw Hill, 2004.
- [2] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz, "A bayesian approach to filtering junk e-mail." In Proceedings of AAAI Workshop on Learning for Text Categorization, pp.55-62, 1998.
- [3] H. Drucker, D. Wu, and V. Vapnik, "Support Vector Machines for Spam Categorization," IEEE Trans. on Neural Networks, Vol.10, No.5, pp.1048-1054, 1999.
- [4] I. Witten and E. Frank, Data Mining: practical machine learning tools and techniques 2nd ed. Morgan Kaufmann, 2005.
- [5] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," Journal of the Royal Statistical Society, Series B, Vol.39, No.1, pp.1 - 38, 1977.
- [6] 김종완, 김희재, 강신재, "데이터 마이닝 기술을 적용한 사용자 선호 스팸 대응 온톨로지 구축," 한국퍼지및지능시스템학회논문지, Vol.17, No.2, pp.160-166, 2007.
- [7] R. Quinlan, C4.5 Programs for Machine Learning, Morgan Kauffmann, 1993.
- [8] T. R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," Int. Journal of Human-Computer Studies Vol.43, pp.907-928, 1995.
- [9] D. Dou, V. McDermott, and P. Qi, "Ontology translation on the semantic web," Journal of Data Semantics, Vol.2, pp.35-57, 2004.